جامعة نايف العربية للعلوم الأمنية
Naif Arab University For Security Sciences

مدى صدق وثبات الاختبارات النفسية في الكشف عن الجريمة في العالم العربي

الدكتور عبدالرحمن عدس

الرياض

1408 هـ - 1988 م

مدى صدق وثبات الاختبارات النفسية في الكشف عن الجريمة في العالم العربي

الدكتور عبدالرحمن عدس(*)

أولا: الصفات المرغوب توفرها في أدوات القياس بشكل عام.

عندما نستخدم اختباراً من أجل الحصول على معلومات تساعدنا في اتخاذ قرار ما فاننا نواجه مشكلة أساسية تتعلق باختيارنا للاختبار الذي يمكن أن يفيدنا حقيقة في اتخاذ القرار، وقد يكون هناك عدة بدائل معروضة لاختبارات يمكن استخدامها لأغراض القرار، ولكن السؤال: أي هذه الاختبارات هو الأفضل في تقديم معلومات مفيدة؟ وما هي الأسس التي نستند اليها في الاختبار؟

هناك أمور كثيرة يمكن أن تؤخذ بعين الاعتبار عند تقويم جودة اختبار ما، ولكننا سنجعلها جميعاً في ثلاثة عناوين رئيسية:

الصدق ـ الثبات ـ القابلية للاستعمال. ويشير:

الصدق: الى الدرجة التي يمكن فيها لاختبار أن يقدم معلومات ذات صلة بالقرار الذي سيبنى عليها، اذ أن الحكم بصدق الاختبار يكون دوماً بدلالة قرار أو استعمال.

الثبات: يشير الى درجة الدقة أو الضبط والاحكام في عملية القياس، وتعطينا معاملات الثبات دلالات عن درجة الاتساق أو التوافق في نتيجة القياس عند تكراره.

(*) عميد كلية الآداب. الجامعة الأردنية. المملكة الأردنية الهاشمية.

القابلية للاستعمال: فتشمل مجموعة من العوامل لها علاقة بالاقتصاد (في الكلفة والوقت والجهد) والملائمة والقابلية للتفسير، وهي عوامل تقرر ما اذا كان استعمال الاختبار على نطاق واسع أمراً عملياً، ويتم تناول هذه الجوانب بشيء من التفصيل في الفقرات التالية:

الصدق: Validity

السؤال الذي يحظى بالأهمية الأولى بالنسبة لأية وسيلة اختبارية هو: كم هي صادقة؟

وعندما نطرح مثل هذا السؤال فاننا في الحقيقة نقضي ما اذا كان الاختبار يقيس ما أردنا له أن يقيس، وليس شيئاً آخر.

عندما نستعمل المتر لقياس طول منضدة لا يكون لدينا أدنى شك في أن المتر يقيس بالفعل طول المنضدة، ويزودنا بمعلومات تفيدنا مباشرة في قرار نريد أن نصل اليه، كأن نقرر فيما اذا كان بالامكان وضع المنضدة بين حافتي نافذتين في الغرفة، ان خبرتنا الطويلة مع هذا النوع من أدوات القياس تؤكد بما لا يدع مجالات للشك صدقها في قياس الطول.

والان لنفرض أننا طبقنا اختباراً في التحصيل القرائي على مجموعة من الطلبة مثل هذا الاختبار يتطلب من الطلبة أن يختاروا اجابات معينة لمجموعة من الأسئلة تتعلق بنصوص في القراءة، وان يضعوا بأقلامهم اشارات معينة على ورقة للاجابة لتدل على الاجابات التي اختارها كل منهم، نعمد بعد ذلك الى تصحيح الأوراق بعد

الاشارات على الاجابات الصحيحة وفق مفتاح للاجابة، معد مسبقاً، ونعطي كل طالب علامة هي عدد الاجابات التي أصاب في اختيارها في ورقته، هنا نعتبر العلامة مقياساً لاستيعاب القرائي عند الطالب، الاَّ أن العلامة نفسها ليست الاستيعاب القرائي وانما هي «سجل لعينة» من السلوك وأية أحكام عن لاستيعاب القرائي انما نستدل عليها من عدد الاجابات الصحيحة، ومن هنا فان صدر العلامة ليس واضحاً للعيان، وانما نتحقق منه خلال البينة المناسبة.

مرة أخرى لنأخذ قائمة للشخصية نحاول من خلالها الحصول على تقدير للتكيف الانفعالي.

عند تطبيق هذه القائمة على مفحوص يطلب اليه أن يؤشر على العبارات التي تصف شعوراً أو سلوكاً يميزه أو لا يميزه.

واستناداً الى اجراءات خاصة فقد حددت استجابات معينة لتدل على سوء التكيف الانفعالي ونحصل على علامة لمفحوص بعد الاستجابات التي اختارها، الاَّ أن الاشارات التي يضعها المفحوص على قطعة من الورق، (القائمة) بعيدة كل البعد عن المظاهر الفعلية للاضطراب الانفعالي، ولابد لنا من ايجاد طريقة نتحقق بها من أن الأداء على الاختبارات يرتبط بدرجة ما بنوع السلوك الذي نهتم به، فكيف يمكننا أن نحدد صدق أداة للقياس مثل هذه؟

أنواع دلالات الصدق:

يمكن النظر الى الاختبار على أنه يمثل جانباً من سلوك الانسان، بمعنى من ثلاثة معانٍ وقد استخدمت المصطلحات التالية للدلالة على هذه المعاني:

١ ـ صدق المحتوى (Content Validity)

٢ ـ والصدق بدلالة محك (Criterion Related Validity).

٣ ـ وصدق المفهوم (Construct Validity).

سنتناول كلا من هذه الأنواع بشيء من التفصيل للوصول الى فهم واضح لما ينطوي عليه كل نوع وكيف تتحدد صلته بكل نوع من الاختبارات.

١ ـ صدق المحتوى:

لنأخذ اختباراً صمم لقياس الكفاية في استعمال اللغة، فكيف يمكننا أن نتعرف الى الدرجة التي يقيس بها هذا الاختبار التحصيل اللغوي؟

أولا: لابد أن نتوصل الى نوع من الاتفاق حول المهارات والمعارف، التي تؤلف الاستعمال السليم والفعال للغة، والتي كانت أهدافها لتدريس اللغة، بعد ذلك لابد لنا أن نتفحص الاختبار نفسه لنرى ما يتطلبه من مهارة ومعرفة وفهم.

وأخيراً لابد لنا أن نجري المطابقة بين محتوى الاختبار وبين تحليل المحتوى للمادة وأهداف تدريسها، لنرى بأية درجة تمثلت

نتائج التحليل هذا في محتوى الاختبار وبالقدر الذي تكون فيها أهداف التدريس المرغوب بها، ممثلة في الاختبار صادقاً لاغراض استعماله في المدرسة.

وبما أن عملية التحليل هي بشكل أساسي عملية منطقية تعتمد على الاجتهاد الشخصي يشار اليها في بعض الأحيان بالصدق المنطقي، وكذلك يستخدم مصطلح المحتوى لأن التحليل يتم على أساس محتوى الاختبار، ومع ذلك يجب أن يؤخذ المحتوى بالمعني الضيق، لأن اهتمامنا يمكن أن يكون بالعمليات بقدر ما هو بالمحتوى البسيط، فقد يكون اهتمامنا في مجال التعبير اللغوي بعناصر في المحتوى من نوع قواعد الاملاء والاعراب والتنقيط، ولكن من جهة أخرى يمكن أن يكون اهتمامنا بعمليات تتعلق بمهارات من نوع تنظيم الأفكار في تسلسل منطقي، أو كتابة جمل تعبر عن وحدة في التفكير أو اختيار الكلمة المناسبة لتدل على المعني المقصود.

وبكلمة أخرى، المحتوى هو ما ينشغل به الطالب والعملية هي ما يفعله بالمحتوى، ان التحقق من صدق المحتوى مسألة ذات صلة وثيقة باعداد جدول المواصفات للاختبار ووضع الاختبار بحيث يناظر جدول المواصفات، ويتحقق صدق المحتوى في الاختبار الذي يضعه المعلم، بالقدر الذي يكون فيه جدول المواصفات حصيلة تحليل واع وذكي لأهداف المادة الدراسية وبالقدر الذي توجه فيه العناية والمهارة والابداع في صياغة فقرات اختبارية تنطبق فيها خصائص جدول المواصفات، ويمكن ان يتحقق لاختبار مقنن صدق

في مدرسة معينة أو منهاج معين، بالقدر يكون فيه محتوى الاختبار مناظراً وممثلا للأهداف المعتمدة في تلك المدرسة أو ذلك المنهاج.

ولابد من التنويه هنا بأن صدق المحتوى أو الصدق المنطقي، له أهميته بالدرجة الأولى في مقاييس التحصيل.

ويهمنا صدق الاختبارات التكوينية (Formative Tests) التي تتناول الكفاية في واحد أو أكثر من الأهداف التربوية الخاصة وكذلك أنواع الاختبارات التي يشار اليها بدلالة المحتوى أو بدلالة (Content -or Criterion-Reference Tests) وفيها يتحقق الصدق بالقدر الذي تكون فيها المهمات في الاختبار ممثلة لأهداف محددة، أما في الاختبارات التجميعية (Summative Tests) واختبارات التحصيل بدلالة معيار المجموعة (Norm-references achievement test) فان اهتمامنا في المقام الأول هو الدرجة التي تكون فيها الاختبارات ممثلة لأحكام الخبراء في أفضل صورها، لما يمكن أن يعتبروه مهماً من معارف ومهارات فاذا جاء التناظر جيداً كان الحكم بأن الاختبار يتمتع بصدق مرتفع واذا جاء التناظر ضعيفاً كان الحكم بأن صدق الاختبار منخفض.

٢ ـ الصدق بدلالة محك:

كثيراً ما يكون اهتمامنا استخدام اختبار لأغراض التوصل الى قرار يتضمن التنبؤ بناتج معين يقع في المستقبل فقد نطبق اختباراً للاستعداد الدراسي للتنبؤ باحتمالات النجاح لخريج المرحلة الثانوية

عند التحاقه بكلية معينة، حيث يعبر عن النجاح ولو بشكل تقريبي بالمعدل الجامعي، كذلك يمكننا أن نطبق اختباراً للتوظيف لانتقاء عمال آلات بحيث يكون الاحتمال كبيراً أن يكونوا موظفين ناجحين.

وهنا يمكن أن نعبر عن النجاح بنوع من المحك، مثل مستوى الانتاج العالي بأقل قدر من الهدر وأدنى نسبة من تاركي العمل، وعندما يكون هذا غرضنا من الاختبار لا نهتم كثيراً بمظهر الاختبار، انما يكون اهتمامنا كله تقريباً بالدرجة التي يرتبط بها مع مقياس للمحك نختاره ليدل على النجاح في العمل، اذن يكون هناك مقياس آخر (غالباً وليس دائماً ما يتشكل في فترة لاحقة في المستقبل) يعتمد محكاً للنجاح ونقدر صدق الاختبار بدرجة العلاقة بينه وبين مقياس المحك، وكلما كان الارتباط مرتفعاً كان الاختبار أكثر صلاحية.

ان تقويم الاختبار من حيث قدرته على التنبؤ عملية تجريبية واحصائية وقد وصف هذا الجانب من الصدق، في بعض الأحيان «بالصدق التجريبي» أو «الصدق الاحصائي».

والاجراء العملي هو أن نعطي الاختبار مجموعة من الأفراد عند التحاقهم بالعمل أو ببرنامج التدريب ومتابعتهم في المستقبل ثم الحصول من كل فرد منهم على قياس في محك محدد للنجاح في العمل أو برنامج التدريب ثم نحسب الارتباط بين العلامات على الاختبارات وقياسات محك النجاح، وكلما كان الارتباط مرتفعاً كان الاختبار أكثر فاعلية في التنبؤ.

مشكلة المحك: قلنا انه يمكن تقدير الصدق التنبؤي باستخراج الارتباط بين علامات الاختبار ومقياس محك مناسب للنجاح في العمل، والعبارة ذات الأهمية الخاصة في هذا القول هي «مقياس محك مناسب».

ومن أصعب المشكلات التي يواجهها الباحث في اختبارات الانتقاء هي العثور على مقياس مقبول للنجاح في العمل لاستعماله كمقياس للمحك في التحقق من صدق الاختبار أو تطويره.

وقد يبدو الأمر في غاية البساطة اذ يقع الاختيار على مقياس من مثل معدل الانتاج أو تقديرات الرؤساء، فيقدر انها تصلح كمقياس محكية وقد يبدو أيضاً أنه ما دام قد اتفق على مقياس المحك فسيكون الوصول اليه أمراً سهلاً ومباشراً، الاَّ أن الأمر لسوء الحظ على غير ما يبدو إذ ان العثور على مقياس معقول للمحك أو تطويره أمر يتطلب في العادة من الباحث في مجال القياس والاختبارات مواجهة عدد من المسائل المعقدة.

ولكل نوع من مقاييس المحك مشكلاته الخاصة به، فمما قد يسترعي الاهتمام سجل الانتاج الفعلي (للفرد) على شكل عدد القطع المصنعة أو الاخطاء في سجلات محاسب أو المبالغ المحصلة في مبيعات وكيل لشركة تأمين، الا أن هناك الكثير من الأعمال مثل طبيب، معلم، وسكرتيرة التي لا يمكن فيها الحصول على تسجيل موضوعي للانتاج أو الانجاز وحتى اذا وجدنا مثل هذه السجلات فكثيراً ما يكون الانتاج متأثراً بعدد من العوامل لا تقع تحت سيطرة

الفرد، فسجل الانتاج لعامل نسيج لا يعتمد فقط على مهارته في نسج الخيوط أو التحكم بالنول، ولكن بالاضافة الى ذلك هناك عوامل أخرى مثل حالة الأجهزة أو الاضاءة في مكان العمل أو لون الخيوط التي ينسجها، كذلك فان مبيعات وكيل التأمين لا تعتمد فقط على كفاءته كبائع ولكن بالاضافة الى ذلك على المنطقة التي يعمل بها ونوع الاشراف والمساعدة اللذين يتلقاهما.

وبسبب عدم وجود سجلات للانجاز أو لأن هذه السجلات في كثير من الأحيان تعاني من جوانب قصور، يلجأ المختص النفسي في شئون الموظفين الى نوع من التقديرات يقدمها الرؤساء كمقياس للمحك، مثل هذه التقديرات يمكن أن تكون جزءاً من الاجراءات الروتينية في المؤسسة، كما هو الأمر عندما نقدم عن كل موظف تقريراً سنوياً أو نصف سنوي، عن كفاءته في العمل، كما يمكن أن تجمع التقديرات لأغراض البحث في نموذج خاص وفي كلا الحالتين سيتبين أن التقديرات تتأثر بالشخص الذي يعطيها والشخص الذي تعمل عنه، وتميل التقديرات الى الخطأ والتأثر بعوامل كثيرة غير تلك المرتبطة بالانجاز في العمل.

وتعتبر مقاييس المحك جميعها جزئية من حيث أنها تقيس جانباً واحداً فقط من النجاح في العمل أو أنها تقيس مجرد أولويات الأداء الفعلي في العمل، وهذا الأمر ينطبق على العلامات في كلية الهندسة مثلا، اذ أنها تمثل محكاً آنياً ونسبياً علاوة على كونها محكاً جزئياً لنجاح المهندس، ان المحك النهائي لابد أن يتناول تقويماً لنجاح الانسان في

مهنته في الحياة اليومية، وبحكم طبيعة الأمور فان هذا المحك النهائي غير متيسر، ولابد للباحث من أن يكتفي ببدائل مناسبة عنه، ومثل هذه البدائل هي في الغالب جزئية، وليست كافية اطلاقاً، وتتلخص مشكلة الباحث عندئذ في اختيار أفضل المقاييس كفاءة، أو أفضل مجموعة مما يبدو له أنه من السهل الحصول عليها، وعندها تكون المشكلة التي يواجهها في القرار حول أي من هذه المقاييس يمكن أن تكون الأكثر فاعلية، فكيف له أن يتخذ مثل هذا القرار؟

الصفات المرغوب فيها في مقياس المحك:

هناك أربع صفات مرغوب فيها في مقياس المحك، وهذه الصفات مرتبة حسب أهميتها هي:

١ ـ الصلة الوثيقة (Relevance) بالموضوع (النجاح في العمل مثلا).

٢ ـ الخلو من التحيز (Freedom of Bias).

٣ ـ الثبات (Reliability).

٤ ـ جدواه أو مدى تيسره (Availability).

ويعتبر المحك وثيق الصلة بالموضوع بالقدر الذي يكون فيه الأداء على مقياس المحك مناظراً أو ممثلا للنجاح في العمل، وفي تقديرنا لهذه الصلة نعود ثانية الى اعتبارات منطقية، فليس هناك بينات تجريبية يمكن أن تدلنا على كم هي الصلة ـ وعلى سبيل المثال ـ بين معدلات السنة الأولى الجامعية والأهداف التي يفترض تحققها في الجامعة النموذجية، (على اعتبار أن العلامات كمقياس للمحك تكشف أو تمثل مدى تحقق الأهداف).

فقد وجد أن اختبارات التحصيل تحتاج الى احكام أفضل المتخصصين لتقرر ما اذا كان محتوى الاختبار يمثل بدقة الأهداف التربوية السليمة، كذلك الأمر بالنسبة لمقياس المحك لابد من الاعتماد على احكام المهني المتخصص لتقدير الدرجة التي يكون مقياس المحك جزئياً وثيق الصلة بالمحك النهائي للنجاح في العمل، أو في برنامج تعليمي.

العامل المهم الثاني في مقياس المحك هو خلوه من التحيز، ويقصد بذلك أن المقياس يجب أن يكون من النوع الذي يتيح لكل فرد نفس الفرصة للحصول على علامة تمثل أحسن ما عنده.

والأمثلة على عوامل التحيز كثيرة منها التفاوت في نوعية الأجهزة وظروف العمل، بالنسبة لعمال المصانع، والتفاوت في «كرم» الرؤساء في وضع تقديرات السكرتيرات أو التفاوت في نوعية التدريس التي يتلقاها الطلبة في صفوف مختلفة، وعندما تعتمد العلامة (في مقياس المحك) على عوامل في ظروف العمل أو في تقويمه أكثر مما تعتمد على الكفاءة في العمل، لا يكون هناك معنى حقيقي للارتباط بين نتائج الاختبار والعلامات على المحك.

أما العامل الثالث وهو الثبات: فستبحث جوانبه العامة في جزء يأتي فيما بعد من هذا البحث، أما من حيث صلته بعلامات المحك فيمكن تلخيصها على النحو التالي: يجب أن يكون مقياس النجاح في العمل مستقراً أي اننا عند تكراره نحصل على نتائج متقاربة جداً، وذلك حتى يمكن التنبؤ به باستخدام أية وسيلة

اختبارية، أما اذا كانت العلامة في المحك تتذبذب بين يوم وآخر والشخص الذي يظهر عنده الانجاز في العمل عالياً هذا الأسبوع يظهر عنده الانجاز منخفضاً في الأسبوع التالي، أو أن الشخص الذي يحصل على تقديرات عالية على عمله من أحد المشرفين يحصل على تقديرات منخفضة من مشرف آخر، عندها يصبح من المستحيل العثور على اختبار يمكن أن يتنبأ بمثل هذه العلامات، عندما يكون المقياس غير مستقر بأي حال لا يمكن التنبؤ بأية طريقة.

وأخيراً فاننا في اختبارنا لمقياس محك معين نواجه دائماً مشكلات تتعلق بدرجة ملائمة المقياس وجدواه ومدى تيسره، كم من الزمن نحتاج للحصول على علامة في المحك لكل فرد؟ وكم هي الكلفة؟ وكم من التعطيل يحدث في النشاط الاعتيادي؟ وبالرغم من أن برامج البحث في شئون العاملين تكون في الغالب ممولة بدرجة كافية بحيث تغطي نفقات الحصول على بيانات جيدة على المحك، الا أن هناك دوماً حدوداً عملية يجب أن تؤخذ بعين الاعتبار عند اختيار مقياس مناسب للمحك.

ان جدوى اختبار التنبؤ لا تعتمد على الدرجة التي يرتبط بها الاختبار مع مقياس المحك فقط ولكن أيضاً على كمية المعلومات الجديدة التي يزودونا بها، من الواضح انه كلما ارتفع الاتباط بين اختبار للتنبؤ ومقياس للمحك كنا أكثر رضا، ولكن هذا معيار نسبي ونحتاج بالاضافة الى ذلك معيار مطلق، كما يجب أن يرتفع معامل الصدق حتى يكون الاختبار ذا نفع؟ وما هو الصدق المقبول؟ وكأننا

نسأل: كم نرتفع الى فوق؟ ومهما كان الأمر سنحاول أن نقدم نوعاً من الاجابة في المؤسسة التي تستخدم اختباراً كأساس في اتخاذ القرار حول ما اذا كانت ستعين أحد المتقدمين للعمل لديها، أو ستقبل طالباً معيناً، فان السؤال الهام، هل يكون تكرار قرار سليم، بالتعيين والقبول، أكثر باستخدام هذا الاختبار مما لو اعتمدنا على محض الصدفة أو لو اعتمدنا على قياس متيسر لنا لكنه أقل صدقاً؟

الاجابة على مثل هذا السؤال تعتمد الى حد بعيد على نسبة الأفراد الذين سنقبلهم ويكون اسلوب الانتقاء أكثر فاعلية فيما لو احتجنا أن نقبل فرداً واحداً، يبدو أنه من الأفضل من بين كل عشرة متقدمين مما لو كانت حاجتنا هي الى تسعة أفراد من كل عشرة، ويبين الجدول (١) النسبة المئوية للاختيارات الصحيحة (أي الناجحين) من بين المقبولين والتي تناظر ارتباطات مختلفة، ويبين الجدول التحسن في «معدل الاصابات» بارتفاع الارتباط، فعندما يكون مقدار الارتباط ٠,٤٠ تكون نسبة الاختيار الصحيح (٦٣,١٪) وعند ارتباط مقداره ٠,٨٠ تصل نسبة الاختيار الصحيح الى ٧٩,٥٪ وهكذا.

الجـــــدول رقم (١)

النسبة المئوية للاختيار الصحيح عندما تكون نسبة القبول ٥٠٪

الصدق لوسيلة الانتقاء	النسبة المئوية للاختيار الصحيح
٠,٠٠	٥٠,٠
٠,٢٠	٥٦,٤
٠,٤٠	٦٣,١
٠,٥٠	٦٦,٧
٠,٦٠	٧٠,٥
٠,٧٠	٧٤,٧
٠,٨٠	٧٩,٥
٠,٩٠	٨٥,٦

ولا يقتصر الجدول على اظهار الدقة المتحققة عند ارتباط معين ولكن مقدار التحسن في الدقة بتحسن الصدق لمقياس التنبؤ، فاذا استطعنا أن نستبدل اختباراً للتنبؤ صدقه ٠,٤٠ بآخر صدقه ٠,٦٠ نكون قد زدنا النسبة المئوية للقرار الصحيح من ٦٣,١ ـ ٧٠,٥، وجميع هذه النسب تشير بالطبع الى القواعد الأساسية المثبتة في الفقرات السابقة، ومع ذلك فان الجدول رقم (١) يقدم نموذجاً مناسباً لفهم المؤثرات في برنامج للانتقاء من وجهة نظر مؤسسات الاستخدام أو الاجازة.

وفي كثير من مواقف الانتقاء يمكن ترجمة التحسن (في القيمة التنبؤية) الى وفر في النفقات مقدراً بقيمة نقدية، فمثلا اذا كانت

شركة تتكبد ٥٠٠ دولار نفقات تشغيل وتدريب عامل جديد حتى يصبح ذا انتاجية مجدية، فان وسيلة الانتقاء التي ترفع النسبة المئوية للناجحين في العمل من ٥٦,٤ الى ٦٣,١ تؤدي الى وفر كلفة التدريب وحده يقدر بـ ٣٣٥٠ دولاراً لكل ١٠٠ مستخدم جديد، ولم يدخل في الحسبان هنا أن العمال الذين تم انتقاؤهم على أساس الاختبار يمكن أن يكونوا أكثر كفاءة بعد انتهاء تدريبهم، وطبعاً لابد من الموازنة بين التوفير الحاصل والزيادة في الكلفة الناشئة من تطبيق وسيلة الانتقاء.

هناك عوامل عدة تميل الى تسوية معاملات الصدق وتعقيد دلالتها، أحد هذه العوامل هو عدم الثبات في مقياس التنبؤ ومقياس المحك المتنبأ به، سيناقش هذا التأثير عند تناول الثبات في جزء لاحق من هذا الفصل.

عامل آخر هو ضيق المدى (Restriction of range) لتوزيع القدرة في المجموعة، بسبب نوع الانتقاء المسبق والثبات المتدني من جهة أو الانتقاء المسبق من جهة أخرى، كلاهما يعمل على تقليص قيم معاملات الصدق.

وهناك نقطة أخرى لابد من التأكيد عليها وهي أن الصدق «نوعي» (Specific) بمعنى أن يتحدد بمنهاج خاص أو وظيفة معينة، وعندما يدعي مؤلف أو ناشر أن اختباره صادق يكون السؤال المناسب دائماً، صادق لأي غرض؟ وبأي معنى؟ فاختبار للتحصيل في الدراسات الاجتماعية الذي تمثلت فيه الأهداف والمحتوى لبرنامج

دراسي معين، قد لا يكون مناسباً تماماً لبرنامج آخر في مكان آخر، ولذلك لابد أن يقوم اختبار التحصيل في ضوء أهداف البرنامج الدراسي المعين، كذلك اذا تبين لنا أن بطارية اختبارات للاستعداد صادقة في انتقاء موظفي مبيعات في محلات تجارية، ممن يتصفون بلطفهم مع الزبائن ملمون جيداً بالسلع التي يتعاملون بها ودقيقون في حساباتهم . . هذه البطارية من الاختبارات يمكن أن تكون عديمة الفائدة في التعرف على وكلاء مبيعات لشركات التأمين، الذين يفترض أن تكون فاعليتهم في الخروج لعمل صفقات التأمين بطرق مختلفة، يجب دائماً أن يقوم صدق المقياس في موقف شبيه بالموقف الذي يستخدم فيه المقياس.

ولكن ما هي أنواع الاختبارات التي يكون فيها الصدق بدلالة محك مهماً، من الواضح أن هذا النوع من الصدق يكون غاية في الأهمية في الاختبارات التي تستخدم في التنبؤ بنواتج يمكن تمثيلها بمقاييس محك، فعندما يكون من السهل العثور على محك للأداء يصلح بشكل جيد لتمثيل النواتج النهائية التي نهتم بها، نكون عندئذ أكثر استعداداً للاعتماد على البينات الناتجة عن العلاقة الارتباطية بين الاختبار ومقياس المحك لترشدنا في قرارنا عن امكانية استخدام الاختبار.

وتتضمن هذه العبارة عنصرين أساسيين: التنبؤ، ومقياس محدد للمحك والمحدد الرئيسي في استخدام الصدق بدلالة محك في سياق التنبؤ يتعلق بدرجة ملائمة المحك المتوفر.

٣ ـ صدق المفهوم:

السؤال الذي نسأله أحياناً عن اختبار نفسي ليس السؤال بأية درجة يتنبأ الاختبار الناجح في العمل؟ وليس السؤال بأية درجة يتمثل المنهاج في الاختبار؟ وانما هو ما معني العلامات المتحققة على الاختبار أو على ماذا تدل؟ ماذا تعطينا العلامة من معلومات عن فرد ما؟

هل تشير الى سمة بارزة أو بناء مما يساعدنا في فهم الفرد؟ ومن أجل أن نقضي ما اذا كان الاختبار يعطينا شيئاً ذا معنى عن الناس فقد استخدم مصطلح صدق البناء (والتعبير الشائع بالعربية صدق المفهوم).

لنأخذ أسلوباً اختبارياً معيناً ولنرَ كيف درس صدقه كمقياس لصفة نفسية ذات معنى أو مفهوم، فقد طور ماكليلاند (McClelland, 1953) أسلوباً اختباريا لتقدير حاجة الفرد أو دافعه الى التحصيل ـ أن يكون الانسان ناجحاً ومجيداً ـ.

وقد استخدمت في الاختبار صور مشابهة لتلك المستخدمة في اختبار تفهم الموضوع (Thematic Apperception Test) يطلب من الفرد في اختبار ماكليلاند أن يؤلف قصة حول كل صورة يروي فيها كيف حدث ما حدث، والى ماذا انتهت الأمور، وقد طور نظام لتصحيح القصص يستند الى احصاء عدد المرات التي تتكرر فيها أفكار عن التفوق والنجاح والتحصيل في مادة القصة، وعلى هذا الأساس حصل كل فرد على علامة تمثل قوة الدافع للتحصيل عنده، والآن كيف لنا أن نقرر ما اذا كان هذا المقياس صادقاً من

حيث أنه يصف بأمانه جانباً ذا معنى في شخصية الفرد؟ لنرى كيف عالج ماكليلاند ومعاونوه هذا الأمر

وقد بدأ هؤلاء الباحثون في الأساس يتساءلون بماذا يجب أن يرتبط مقياس الدافع للتحصيل؟ ثم وضعوا سلسلة من التنبؤات كان بعضها ما يلي:

١ ـ الأفراد الذين يكون الدافع للتحصيل عندهم مرتفعاً يكون أداؤهم في الكلية جيداً تبعاً لاستعدادهم الدراسي.
٢ ـ يكون الدافع للتحصيل مرتفعاً عند الطلبة في الفترة التي انتهوا فيها مباشرة من الاجابة على اختبارات وصفت لهم بأنها تقيس ذكاءهم.
٣ ـ الافراد الذين يكون الدافع للتحصيل عندهم مرتفعاً يكملون عدداً أكبر من الفقرات في اختبار سرعة مشوق.
٤ ـ يكون الدافع للتحصيل أعلى عند الأطفال الذين يأتون من أسر تؤكد على الاستقلالية المبكرة.

لقد استند كل من هذه التنبؤات الى نوع من «النظرية في السلوك البشري» وعلى أساسها نظر الى التحصيل الأكاديمي على أنه تأليف بين القدرة والجهد، وبالتالي فان أولئك الذين يكون دافعهم للتحصيل أعلى يبذلون جهداً وعزماً أكبر فاذا تساوت قدراتهم مع غيرهم كانت علاماتهم الأعلى، وباستخدام تسلسل منطقي مماثل، يمكن الاستدلال على التنبؤات الأخرى، وعلى العموم فقد وجد (ماكليلاند) أن معظم تنبؤاته قد تدعمت بالنتائج التجريبية،

والحقيقة التي تبينت عندما ارتبطت العلامات على الاختبار بعدد من الوقائع بالطريقة التي تم التنبؤ بها من التحليل المنطقي للسمة التي يفترض ان الاختبار يقيسها دعت صدق الاختبار باعتباره يقيس «مفهوماً» أو سمة ذات معنى أمكن تلخيص خصائصها الأساسية تحت اسم «دافع التحصيل».

وهناك عدد كبير من الاختبارات النفسية وعدد أقل من الاختبارات التربوية المعدة لقياس سمات أو صفات عامة في الفرد، ومن أمثلتها: الاستدلال اللفظي Verbal Reasoning والتصور المكاني Spatial Visualizing والقابلية الاجتماعية Sociability والانطواء Introversion والميل الميكانيكي Mechanical Interest وجميعها صفات أو مفاهيم، وتعتبر اختبارات مثل هذه الوظائف صادقة بالقدر الذي تظهر فيه الاختبارات دلالات مناظرة للطريقة التي يتوقع أن تعبر بها هذه الوظائف عن نفسها.

ان نظرية عن السمة تؤدي الى عمل تنبؤات من الأنواع التالية والتي يمكن اختبارها للتأكد من صحتها:

١ ـ تنبؤات عن ارتباطات:

ان طبيعة السمة وبالتالي المقاييس الصادقة التي تقيسها تشير الى أنها لابد من أن ترتبط بمقاييس معينة أخرى، هذه المقاييس الأخرى يمكن أن تكون مقاييس معروفة عن الوظيفة ذاتها، من ذلك أن كثيراً من مقاييس الذكاء الجمعية اللاحقة جرى التحقيق من

صدقها باستخراج ارتباط من مقاييس سابقة، وبخاصة ستانفورد ـ بينية (Stanford-Binet) الذي يطبق كاختبار فردي، من جهة أخرى يمكن أن تكون المقاييس الأخرى من النوع الذي ترتبط به السمة ارتباطاً منطقياً، وعلى هذا الأساس فقد استخرج بعض دلالات الصدق لاختبارات الذكاء باستخراج ارتباطها مع الأداء المدرسي، ولمقاييس الاستعداد الميكانيكي باستخراج ارتباطها مع الكفاية في أعمال ميكانيكية.

ومن الطرق لدراسة المفاهيم أو الأبعاد التي يقيسها اختبار دراسة الارتباطات الداخلية (intercorrelations) بين الاختبار وعدد من الاختبارات الأخرى، والنمط الذي يتشكل من هذه الارتباطات يجعل من الممكن التعرف على الاختبارات التي تشترك في قياس عامل واحد، وبتفحص الاختبارات التي تنجح مع بعضها البعض في عامل واحد نتمكن من جلاء طبيعة العامل ومعناه والاختبارات التي تقيسه، على أن هذا النوع من الصدق «الداخلي» أو «العاملي» لازال بحاجة الى نوع من البينة عن علاقته بوقائع الحياة، خارج الاختبارات نفسها، اذا أردنا أن يكون للعامل أي دلالة واقعية.

وقد يكون من المناسب أحياناً ان نتنبأ أن مقياساً للمفهوم (س) سوف يظهر ارتباطاً منخفضاً أو معدوماً مع مقاييس لابعاد أخرى، أي أن المفهوم مختلف (مستقل) عن هذه الأبعاد، مثال ذلك أن اختباراً مصمماً لقياس الاستيعاب الميكانيكي (Mechanical comprehension) يفترض أن يظهر ارتباطاً منخفضاً نوعاً ما مع

اختبار للقدرة اللفظية (Verbal ability) كذلك يفترض في مقياس للقابلية الاجتماعي (Sociability) أن يرتبط ارتباطاً عالياً جداً مع مقياس للاصرار (assertiveness) اذا كان الاختباران أصلا يقيسان مفهومين مختلفين.

٢ ـ تنبؤات عن الفروق بين الجماعات:

قد تتضمن النظرية افتراضاً بأن مجموعات معينة يكون أداؤها عالياً أو منخفضاً بشكل متميز في سمة معينة، وبالتالي في الاختبار الذي يقيسها وهكذا، قد يبدو معقولا أن يكون أداء مجموعة من البائعين عالياً على مقياس للسيادة (ascendance) أو للاصرار (assertiveness) بينما يكون أداء أمناء المكتبات ـ على المقياس نفسه ـ منخفضاً، وقد نتنبأ أن الأطفال لآباء مهنيين أو من رجال الأعمال هم أكثر سيادة من الأطفال لآباء من المكتبة أو الحرفيين، واذا أخذنا سمات معينة يمكننا من معرفتنا للمجتمع وللفئات التي يتألف منها أن نقدر أنماطاً من الفروق بين هذه الفئات قد يكون لها ما يبررها، وعندما يطبق الباحث اختباراته على هذه الفئات يمكن أن يكتشف ما اذا كانت تنبؤاته ستتحقق.

٣ ـ تنبؤات عن الاستجابة للمعالجة التجريبية experimental treatment أو التدخل (intervention)

نستدل من النظرية أن الطريقة التي يعبر بها الانسان عن خصائصه يمكن تعديلها تحت تأثير ظروف أو معالجات تجريبية، وهكذا قد يكون من المعقول أن نتنبأ بأن القلق عند شخص يزداد في

الفترة السابقة مباشرة لاجراء عملية جراحية ثانوية عليه، وقد استخدم ما يسمى «بمعدل اندماج الومضات المتناوبة» (Rate of Flicker Fusion) كمقياس لمستوى القلق، ففي احدى الدراسات (Buhler, 1953) كانت النتيجة ـ التي سبق التنبؤ بها ـ ان عتبة الاندماج للومضات المتناوبة كان قبل اجراء العملية أدني مما هو بعدها، حيث يفترض ان حالة القلق انتهت.

ان أي اختبار يفترض فيه أن يقيس سمة أو وظيفة، باستطاعتنا ان نصوغ له اطاراً نظرياً يمكننا من اشتقاق تنبؤات محددة، هذه التنبؤات يمكن اختبارها، وبالقدر الذي تتحقق فيه يتأكد صدق الاختبار كمقياس للسمة أو المفهوم، فاذا لم تتحقق التنبؤات بالتجربة، ترتب على ذلك ان نتشكك بصدق الاختبار، أو بالنظرية أو بكليهما، ان بينة الصدق هي كما رأينا في جزء منها منطقية وفي جزء آخر تجريبية، والاعتبارات المنطقية لما نقيسه تكون في محور اهتمامنا عندما يكون اهتمامنا بناتج نهائي، يمكن أن يكون اختباراً يستخدم لوصف التعلم السابق لفرد، أو يكون نوعاً من المؤشر نعتمده كمقياس للمحك، كذلك فان احكام نظرية يمكن بواسطتها اختبار صدق مقياس عن مفهوم نفسي هو ممارسة منطقية، وفي المقابل فاننا عندما سنستخدم اختباراً كوسيلة انتقاء، للتنبؤ بنوع من المحك المقبول، يصبح التحقق من الصدق مسألة احصائية، كذلك فان التحقق من النظرية التي نبنيها حول اختبار لمفهوم معين هو أيضاً مسألة تجريبية، الحكم والبينة يلتقيان معاً في خطة التحقق من الصدق.

السؤال الثاني الذي نطرحه عن وسيلة القياس هو: كم هي ثابتة؟ ونحن الآن لا نسأل ماذا تقيسه الوسيلة؟ وانما ما الدقة التي تقيس بها أي شيء تقيسه، وما درجة الضبط والأحكام في العلاقة الناتجة؟ وكم يتحقق فيها من الدقة عند تكرر القياس على الفرد نفسه؟ افرض أننا سنختبر جميع الأولاد في أحد صفوف المدرسة اليوم، ومرة ثانية في الغد لكي نعرف المسافة التي يستطيع كل منهم أن يقذف بالكره اليها، نرسم خط الابتداء على الملعب، ونعطي كل ولد احدى الكرات القديمة التي تحتفظ بها عادة دائرة التربية الرياضية لاغراض التدريب، ونرسل الى الجهة المقابلة احد المراقبين لكي يضع علامة في المكان الذي تصيبه الكره المقذوفة، ونطلب من كل ولد أن يقذف الكرة الى أقصى ما يستطيع، نأتي بشريط متري ونقيس المسافة من خط الابتداء حتى العلامة التي ارتطمت عندها الكرة بالأرض. نتيح لكل ولد أن يقذف الكرة مرة واحدة هذا اليوم ومرة ثانية الغد، فنسجل لكل ولد قياسين.

عندما نقارن للولد الواحد القياسين سنجدهما عند معظم الأولاد لا يتطابقان تماماً، هناك تفاوت في مدى الضربتين، سنجد ان غالبية الفروق صغيرة نسبياً ولكن بعض هذه الفروق كبيرة الى حد ما وقليل منها كبير جداً، ويتبين لنا من هذه الفروق ان المحاولة الواحدة لم تكن كافية لتعطينا مقياساً ثابتاً تماماً لقدرة الولد في قذف الكرة، مثل هذه النتائج تبدو الى حد ما غير متوافقة بين يوم وآخر، فلماذا؟

يمكننا التعرف الى ثلاث فئات عن الأسباب لعدم التوافق بين محاولة اليوم ومحاولة الغد.

١ ـ يمكن أن يكون الشخص قد تغير فعلا بين اليوم الأول واليوم الثاني، ربما كان في اليوم الأول مستريحا أكثر مما كان في الثاني، وربما في أحد اليومين أكثر تشوقاً لمحاولة أفضل ما عنده، وربما حصل على بعض التوجيهات من أبيه فيما بين المحاولتين، ولو كانت الفترة بين المحاولتين بالأشهر وليست أياماً فلربما حدث نمو حقيقي في الجسم، ولكن ليس بنفس الدرجة عند جميع الأولاد، ومثالنا هذا يتناول تغيرات ممكنه في الأداء الجسمي، ويصبح من السهل علينا ان نفكر بتغيرات مماثلة تنطبق في حالة اختبار في القدرة العقلية أو قائمة للتقرير الذاتي (self-report inventory) تتعلق بالمزاج أو الميل.

٢ ـ يمكن أن تكون المهمه اختلفت بين القياسين، مثال ذلك أن الكرة ربما كانت في يوم منفوخة جيداً، وفي اليوم الآخر رخوة بحيث لا يتمكن صاحبها منها جيداً أو ربما كان الفاحص يسمح للاولاد في اليوم الاول أن يركضوا مسافة حتي خط الانطلاق، بينما لم يسمح لهم في اليوم الآخر أكثر من خطوتين، ومثل هذا التفاوت يمكن أن يكون بعض الأولاد استفاد منه أكثر من البعض الآخر، وفي الاختبارات التحريرية كثيراً ما نستخدم صورة من الاختبار في جلسة معينة، ونسخدم صورة أخرى مكافئة للأولى في الجلسة الثانية، في مثل هذه الحالة تكون فقرات الصورتين مختلفة ويمكن

أن يكون بعض الطلبة أحسن حالا في تعاملهم مع عينة من المهمات بينما البعض الآخر أحسن حالا في العينة الثانية.

٣ ـ عينة السلوك المحدودة يترتب عليها أن العلامة غير مستقرة ولا يعول عليها، ولو أننا سمحنا لكل ولد بمحاولتين بنفس الكرة، ونفس التعليمات وبفاصل للاستراحه بين المحاولتين من خمس دقائق فقط فان مدى الرميتين نادراً ما يأتي نفسه تماماً فالمحاولة المفردة هي عينة هزيلة من السلوك، مثل هذه العينة وتقويمها معرضان لتأثيرات الصدفة بأشكال كثيرة مختلفة، وربما انزلق أصبع، الولد ربما اضطرب التآزر بين ساقيه وذراعه . . أو ربما أمسك بالكرة بعيدة أو قريبة منه أكثر مما يجب . . ربما كان الشخص الذي سجل الهدف ينظر الى ناحية أخرى عندما ارتطمت الكرة بالأرض، ربما جاءت هبة ريح في اللحظة التي ألقى فيها بالكرة، أو ربما حدث أمر من مائة، بعضها في صالحه وبعضها في غير صالحه، كل هذه الأمور تقودنا الى نتيجة واحدة وهي أن العينة الصغيرة من السلوك لا تعطينا وصفاً للفرد ثابتاً يمكن الوثوق به، سواء كانت العينة قذف كرات الى مسافة معينة أو قراءة جمل لاغراض الفهم.

هناك طريقتان للتعبير عن الثبات أو الدقة في مجموعة من القياسات، أو من وجهة نظر عكسية عن التفاوت (التباين) داخل المجموعة، الطريقة الأولى تشير بشكل مباشر الى كمية التباين المتوقعة في مجموعة من القياسات المتكررة على فرد واحد، فلو أمكن أن نجعل

خالداً يلقي بالكرة ٢٠٠ مرة على (افتراض أن هذا الأمر ممكن دون أن تتدخل مؤثرات الارهاق والتدريب) لحصلنا على توزيع تكراري للمسافات التي القيت اليها الكرة، ولهذا التوزيع التكراري متوسط يمكن اعتبار أنه يقترب من القيمة الحقيقية للمسافة التي يستطيع خالد أن يلقي بالكرة اليها، وله أيضاً انحراف معياري يصف تشتت هذه المسافات، وسنطلق على مقياس التشتت هذا الخطأ المعياري في المقياس (Standard error of measurement) لأن هذا هو الانحراف المعياري «للأخطاء» في قياس مسافة الرمية لشخص واحد.

في البيانات النفسية ولا يتيسر لنا عادة أن نجري سلسلة متكاملة من القياس على كل فرد، فهناك تأثيرات التدريب والتعب، بالاضافة الى ذلك لا يتسع لنا الوقت أن نطبق ٢٠٠ اختبار للقراءة أو ٢٠٠ قائمة للميول، في بعض الأحيان يحالفنا الحظ بالحصول على علامتين اثنتين لكل فرد، ولكن اذا حصلنا على قياسين لكل فرد، نستطيع أن نتوصل الى تقدير لتشتت العلامات التي يفترض أن يحصل عليها الشخص المتوسط فيما لو تكرر قياسه عدداً كبيراً من المرات.

القياسات الثابتة تعني ان الفرد يحافظ على نفس الموقع تقريباً بالنسبة لمجموعته، فالولد الذي حصل على أعلى قيمة للمسافة في اختبار قذف الكرة أول مرة، يجب أن يكون هو نفسه الذي يحصل على قيمة في المرة التالية، وكذلك يحافظ على كل شخص في المجموعة

على موقعه نفسه تقريباً، اننا نعلم أن معامل الارتباط يزودنا بدلالة احصائية للدرجة التي يتواكب أو يتلازم فيها معاً شيئان، المرتفع مع المرتفع والمنخفض مع المنخفض، فاذا كان الشيئان المترابطان هما تطبيقان لمقياس واحد فان معامل الارتباط الناتج يعطينا مؤشراً للثبات، ونستطيع أن نطلق عليه معامل الثبات (reliability coefficient) وخصائص معامل الارتباط هذا هي نفسها التي سبق الكلام عنها في مناقشاتنا، ولكن العلاقة التي أمامنا الأن هي من مجموعتين من القياسات بنفس المقياس وكلما كان الأفراد مرتبين بنفس الرتب في المرتين، كان الارتباط أعلى والاختبار أكثر ثباتاً.

اذن يكون المقياس ثابتاً بالقدر الذي يبقى فيه الفرد على حاله تقريباً عند تكرار قياسه، ويبقى على حاله تقريباً بالقدر الذي يتمثل فيه بقيمة صغيرة للخطأ المعياري في القياس، أو بمعامل ثبات مرتفع ولكن ما نوع البيانات على وجه التحديد، التي نحتاجها حتى نحصل على تقدير معقول لدرجة الاستقرار أو الدقة في القياس؟

سوف نتناول ثلاث طرق مميزة ونلاحظ أوجه النسبة والاختلاف فيما بينها كما تقوِّم حسنات وسيئات كل منها:

١ ـ تكرار نفس الاختبار أو القياس.
٢ ـ تطبيق صورة مكافئة ثابتة من الاختبار.
٣ ـ تجزئة الاختبار الى جزءين أو أكثر من الأجزاء المتكافئة.

فلنتمعن في كل منها على الترتيب:

١ ـ اعادة الاختبار نفسه:

اذا كنا نرغب في أن نتعرف الى درجة الثبات التي نقوّم بها قذف الكرة عند فرد فيمكننا أن نختبره مرتين، وقد نرى من المناسب أن نتخذ جانب الحيطة بان نطلب من شخصين مختلفين أن يسجلا القياسين كل منهما بشكل مستقل عن الآخر.

اذ لا نريد للمجرب أن يتأثر تسجيله للقياس الثاني بما يتذكره عن القياس الأول، وقد يكون من المرغوب فيه أن نجري الاختبارين في يومين مختلفين، وهذا يعتمد على محور اهتمامنا فاذا كنا نرغب في أن نتعرف على الدقة التي يمكن لقذفة واحدة (ويمكن أن تكون مجموعة من القذفات) أن تصف بها شخصاً ما في نقطة زمنية معينة ترتب على ذلك أن نجري القياسين واحداً بعد الآخر وعندها نعلم أن الشخص بقي على حاله، وان مصدر التباين أو الخطأ يمكن في عمليه قياسه، أما اذا كنا نرغب في ان نعرف الدقة التي يمكن لقياس معين ان يصف بها شخصاً ما بين يوم وآخر، أي الدرجة التي نتنبأ بها بأن ما يسجله في الأسبوع التالي يقترب مما يسجله هذا اليوم، عندها يكون من المناسب ان نجري عملية القياس في مناسبتين منفصلتين، هنا يكون اهتمامنا في تباين الفرد من وقت الى آخر كما هي في التباين العائد الى عملية القياس.

وفي بعض الأحيان يكون اهتمامنا بالتباين في الفرد من يوم الى آخر وفي بعض الأحيان لا نهتم بذلك، هنا يمكن أن نتساءل: ما الدقة التي تصف بها الفرد «م» لقياسنا له في هذه اللحظة؟ أو قد

نتساءل ما الدقة التي يمكن أن بقياسنا للفرد «م» اليوم ان تصف بها ما يكون عليه في الغد أو بعد اسبوع، أو بعد شهر.

كلا السؤالين له ما يبرره، ولكنهما ليسا نفس السؤال فالبيانات التي نحتاج لجمعها للاجابة عن السؤالين تختلف عن البيانات التي نحتاج للاجابة عن الآخر.

ولكي ندرس ثبات القياس في صفة فيزيائية للشخص كطوله أو وزنه فان تكرار القياس عملية مباشرة وكافية، ويبدو انها أيضا كافية وقابلة للتطبيق في بعض جوانب السلوك مثل سرعة الاستجابة أو أنواع المهارات الحركية الممثلة بقذف الكرة، ولكن افرض ان اهتمامنا الآن هو في ثبات اختبار للاستيعاب القرائي ولنفرض أن الاختبار مؤلف من ستة نصوص، لكل منها خمسة أسئلة. ولنفرض ان الاختبار طبق مرة، وبعد ذلك مباشرة طبق مرة ثانية، فماذا يحدث؟ من المؤكد ان المفحوص لا يحتاج أن يعيد قراءة المادة التي سبق له أن قرأها قبل لحظات.

ربما قرأ بعضها ولكن اجاباته في المرة الثانية ستعتمد الى حد كبير على مجرد تذكر الاجابات التي اختارها في المرة الأولى، فيسجلها ثانية، واذا لم يكن بمقدوره أن ينهي الاختبار في المرة الأولى، فسيتمكن هذه المرة من التقدم واستغلال معظم وقته في المادة الجديدة، ان مثل هذه التأثيرات تظل بدرجة ما سارية المفعول حتى لو كانت الفترة (بين مرتي التطبيق) طويلة، ومن الواضح ان اختباراً

مثل هذا في المرة الثانية لا يمثل المهمات نفسها تماماً كما عرضت في المرة الأولى.

هناك اعتبار آخر يظهر عند تكرار اختبار من نوع اختبار الاستيعاب القرائي، فلنفرض ان احد النصوص الخمسة في الاختبار كان عن كرة القدم، وان احد المفحوصين كان خبيراً في كرة القدم، سيكون هذا الفحص، بالطبع سهلاً بالنسبة له، ومن شأن هذا ان يميزه عن غيره ببعض النقاط، ويمكن أن يعمل مثل هذا الاختبار على رفع تقدير مستواه القرائي عن تقديره الحقيقي، ونلاحظ هنا مثل هذا التقدير يتكرر في مرتي التطبيق اذا بقيت مادة الاختبار نفسها، ويظل الخطأ ـ للفرد «م» خطأ ثابتاً (Constant error) في مرتي التطبيق، وبما أن هذا الخطأ يؤثر على كلا العلامتين بنفس الطريقة، فانه يجعل الاختبار «يظهر» وكأنه ثابت ـ بدلا من أن يسهم في عدم ثباته، وفي مجال للقدرة مثل القراءة، لابد ان نلاحظ ان أداء فرد قد لا يكون منسقا تماما في جميع جوانب المجال، اذ ان اهتماماته الخاصة، وخبراته وخلفيته تجعل عنده جوانب قوة وجوانب ضعف، فالاختبار المعين هو عينه من المجال بأكمله ومستوى الأداء الذي يحصله الفرد «م» على الاختبار بالمقارنة مع أداء غيره، يعتمد بدرجة ما على العينة الخاصة من المهمات التي تم اختبارها من منطقة (domain) القدرة أو الشخصية التي نحاول تقويمها، اذا ظلت العينة نفسها في كلا القياسين، واستمر سلوكه على حاله بدرجة أكبر مما لو اختلفت عينة المهمات.

لاحظ أنه حتى هذه اللحظة يمكن التعرف على ثلاثة مصادر رئيسية للتباين في الأداء يمكن ان تقلل من الدقة في علامة معينة عندما نحاول أن نصف بها فرداً ما:

١ ـ التباين من محاولة الى أخرى عند الاستجابة الى مهمة معينة في لحظة زمنية معينة.

٢ ـ التباين في الفرد من وقت الى آخر.

٣ ـ التباين الناشيء من انتقاء عينة خاصة من المهمات لتمثل منطقة من السلوك.

ان اعادة الاختبار نفسه على الفرد يمكن أن يتم بطرق تؤدي الى النوعين الأولين من «الخطأ» ولكن لا يمكن بواسطتها تقويم النوع الثالث، أضف الى ذلك تأثيرات ممكنة ناشئة عن التذكر والمران، اللذين سبقت الاشارة اليهما أعلاه.

١ ـ الصور المتكافئة للاختبار Parallel Test Forms.

ان الاهتمام بالمصدر الثالث من التباين وهو التباين العائد الى العينة الخاصة من المهمات المختارة لتمثيل منطقة سلوك معينة، يتطلب مجموعة أخرى من الاجراءات لتقويم الثبات، فاذا كانت عينة الفقرات مصدراً للخطأ ذا دلالة واذا كان ما نجري عليه عادة هو أن نتقصى درجة الدقة التي يمكن أن نعمم بها من علامة محددة تستند الى عينة من المهمات ومن الممكن تحقيق ذلك باستخراج الارتباط بين صورتين متكافئتين من الاختبار.

تعتبر الصور المتكافئة للاختبار صوراً أو أشكالاً بنيت طبقاً لمواصفات واحدة ولكنها تألفت من عينات مستقلة من منطقة سلوك محددة، وعلى هذا فان اختبارين متكافئين للقراءة يجب أن يتضمنا نصوصاً وأسئلة لها نفس الصعوبة، ويسأل فيهما نفس النوع من الأسئلة، أي أن التوازن واحد بين الأسئلة التي تتناول حقائق خاصة وتلك التي تتناول أفكاراً عامة.

اذا كان لدينا صورتان من الاختبار فيمكننا أن نطبق على كل تلميذ احدى الصورتين ثم نتبعها بالأخرى ويمكن أن تتبع الواحدة الثانية مباشرة، اذا لم يكن لنا اهتمام بالاستقرار عبر الزمن، أو نجعل بينها فترة من الزمن اذا كنا نهتم بدرجة الاستقرار، وبحساب الارتباط بين الصورتين نحصل على معامل مناسب للثبات، واذا سمحنا بفترة بين تطبيق الصورتين فان مصادر التباين الثلاثة يكون لها الفرصة أن تظهر تأثيراتها أي التباين الناشيء عن عملية القياس ذاتها والتباين في الفرد من وقت الى آخر والتباين العائد الى عينة المهمات.

ان مطلبنا بأن تكون نتائج الاختبار متسقة تحت مثل هذه الشروط هو المعيار الصارم الذي يمكننا أن نضعه له، فاذا كنا نريد ان نستخدم نتائج الاختبار لاصدار تعميم حول ما يمكن لخالد أن ينجزه في مهمات أخرى، لها نفس الطابع العام، في الأسبوع القادم وفي الشهر القادم فهذا هو المعيار المناسب الذي يمكن به تقويم الاختبار، وهذه هي الطريقة التي نود ان نستخدم لها نتائج الاختبارات في معظم المواقف التربوية، ومن هنا فان البينة المستخلصة من صور

متكافئة للاختبار يجب ان تعطي الوزن الأكبر في تقويم ثبات الاختبار.

ان طريقة الصور المتكافئة للاختبار تقدم لنا أساساً سليماً جداً لتقدير الدقة في اختبار نفسي أو تربوي الاَّ أن هذه الطريقة تثير عدداً من المشكلات العلمية، فهي تتطلب توفر صورتين متكافئتين للاختبار، وتوفر وقتاً يسمح باختبار كل فرد مرتين، في بعض الأحيان لا توجد صورة ثانية من الاختبار أو قد لا يتوفر الوقت لاجراء تطبيق ثانٍ، فتطبيق اختبار ثانٍ منفصل كثيراً ما يمثل مطلباً ثقيلا على الموارد المتاحة، مثل هذه الاعتبارات العلمية المتعلقة بتوافر الامكانات جعلت الكثيرين من مؤلفي الاختبارات يتوجهون الى اجراءات تعتمد على استخلاص تقدير للثبات من تطبيق صورة واحدة فقط من الاختبار.

علماً أن مثل هذه الأساليب هي في أحسن الأحوال مجرد حلول وسط، والطريقة التي تعتمد على استخراج الارتباط بين صورتين مكافئتين، تطبقان في العادة بفاصل زمني يمتد الى عدة أيام أو عدة أسابيع، وهذه تمثل الطريقة المفضلة في تقدير الثبات.

تجزئة الاختبار Sub-divided Test

كثيراً ما تفرض علينا اعتبارات عملية أن نحصل على تقدير للثبات من تطبيق الاختبار مرة واحدة، من الطرق الشائعة الاستعمال لهذا الغرض تقسيم الاختبار الى نصفين، يفترض أنهما

متكافئان ومن الممكن تجميع نصفي الاختبار على أساس تفحص دقيق للمحتوى والصعوبة لكل فقرة وبذل جهد منظم لموازنة المحتوى ومستوى الصعوبة في النصفين، لكن الطريقة الأبسط هذه والتي يكثر استخدامها في تشكيل نصفين مكافئين هي أخذ الفقرات في نصفي الاختبار على التناوب، وذلك بوضع جميع الفقرات ذات الأرقام الفردية في أحد النصفين، والفقرات ذات الارقام الزوجية في النصف الاخر، ويعتبر هذا الأسلوب مناسباً ما دام الاحتمال كبيراً بين الفقرات تتشابه في الشكل والمحتوى والصعوبة، تتجمع مع بعضها بعضاً في الاختبار الواحد، ففي اختبار طويل نسبياً يصل الى ٦٠ فقرة أو أكثر تميل التجزئة بهذه الطريقة الى تحقيق الموازنه في عوامل من نوع شكل الفقرة وشمول محتواها ومستوى صعوبتها وتكون الاحتمالات حسنة أن يكون نصفا الاختبار المتشكلان بهذه الطريقة «متكافئين» وفق الافتراضات التي عرف بها التكافؤ في الجزء السابق.

ويجب ان نتذكر ان الطرق التي نعرضها هنا في تجزئة الاختبار هي فقط لاغراض تصحيحه وليس تطبيقه، أي أن الاختبار يعطي بكامله في جلسة واحدة وفي حدود زمنية واحدة، الّا انه بعد التطبيق تشتق علامتان منفصلتان، واحده من تصحيح الفقرات ذات الارقام الفردية، والأخرى من تصحيح الفقرات ذات الأرقام الزوجية، ومعامل الارتباط بين هاتين العلامتين يعطينا مقياساً للدقة التي يقيس بها الاختبار الفردُ.

من ناحية ثانية يجب أن نلاحظ أن الارتباط المحسوب هو بين نصفي طول الاختبار، وهذه القيمة لا تنطبق مباشرة على الاختبار

بكامل طوله، وهو أداة القياس المعدة للاستعمال وعلى العموم كلما كبرت عينة سلوك الفرد لدينا كان قياسها أكثر ثباتاً، وكلما سجلنا من السلوك أكثر كلما قل اعتماد القياس على عناصر الصدق في سلوك الفرد أو على العينة الخاصة للمهمات اذ تتعادل تقريبا الاجابات الصحيحة التي جاءت بالحظ والاجابات الخطأ التي جاءت نتيجة شرود الذهن أو عدم الانتباه.

عندما يكون نصفا الاختبار اللذان حسب الارتباط بين علامتيهما، متكافئين يمكن الحصول على تقدير غير متحيز (Unbiased estimate) لثبات الاختبار بكامل طوله من الارتباط بين النصفين، وذلك باستخدام المعادلة التالية:

$$\text{ر}_{11} = \frac{2 \cdot \text{ر}_{\frac{1}{2}\frac{1}{2}}}{1 + \text{ر}_{\frac{1}{2}\frac{1}{2}}}$$

حيث ر١١ هي القيمة المقدرة لثبات الاختبار بكامل طوله و $\text{ر}_{\frac{1}{2}\frac{1}{2}}$ الارتباط الفعلي بين نصفي الاختبار.

فاذا فرضنا أن الارتباط بين نصفي اختبار هو ٠,٦٠ فان تطبيق المعادلة (١) يعطينا:

$$\text{ر}_{11} = \frac{2 \times (٠,٦٠)}{1 + ٠,٦٠} = \frac{1,2}{1,6} = ٠,٧٥$$

هذه المعادلة التي يشار اليها عموماً بمعادلة سبيرمان ـ بروان التنبؤية (Spearman-Brown Prophecy Formulae) تبعاً لوظيفتها واسماء واضعيها تجعل من الممكن حساب تقدير للثبات من مجرد تطبيق واحد للاختبار.

ولسهولة تناول طريقة التجزئة النصفية راقت للكثيرين وشاع استعمالها على نطاق واسع، ولذلك نجد أن كثيراً من أدلة الاختبارات (المنشورة) تنقل هذا النوع من معامل الثبات وليس غيره ولسوء الحظ فان لهذا المعامل عدداً من المحدودات (جوانب قصور) لابد من تناولها.

أولا: لما كنا نستخلص العلامتين من تطبيق واحد، فلابد أن العلامتين كلاهما تمثلان الفرد كما هو حاله في لحظة التطبيق بذاتها، وحتى الوقائع التي لا تدوم الا لدقائق معدودة فانها تؤثر على العلامتين بنفس الدرجة، وبكلمة أخرى فان التباين في الفرد من يوم لآخر لا يكشفه هذا النوع من معامل الثبات، ولكنه يمكن فقط أن يعطي دلالة على الدقة التي يقوم بها الفرد في لحظة معينة من الزمن..

عامل آخر يجعل نصفي الاختبار في بعض الأحيان أكثر تشابها مما يتحقق في الصور المتكافئة المنفصلة، فاذا كان الاختبار يتضمن مجموعات من الفقرات تتعلق بمادة مرجعية مشتركة فقرات القراءة المتعلقة بنفس النص، أو فقرات العلوم المتعلقة بنفس التجربة، وغيرها فان الأداء على جميع هذه الفقرات سيعتمد الى حد ما، على فعل مشترك يرتبط بفهم المادة المرجعية ويكون عندئذ الاحتمال كبيراً

بأن المفحوص الذي ينجح في فقرة من المجموعة سينجح في باقي الفقرات.

فالفقرات ليست مستقلة تجريبياً، واذا ما قسمنا مثل هذه المجموعات بين نصفي الاختبار، كما هو الحال في حساب معامل الثبات بتقسيم الفقرات الى فردية وزوجية (Odd-even reliability) يتضخم التشابه بين النصفين بدرجة غير معقولة، وعندما تكون مادة الاختبار من هذا النوع فالاجراء المفضل هو توزيع الفقرات في مجموعات على التناوب بين نصفي الاختبار، اذا كنا نريد معامل الارتباط بين نصفي الاختبار ان يعطي تقديراً معقولا لثبات الاختبار بكامل طوله، تطلب ذلك ان يكون النصفان مستقلين تجريبياً ومتكافئين بمعنى أنه لا يوجد تداخل لا مبرر له بين الفقرات في أحد النصفين والفقرات في النصف الآخر، وفي نفس الوقت بين كل من النصفين بحيث تتحقق فيه المواصفات نفسها للاختبار بكليته.

ثانياً: معامل الثبات بالتجزئة النصفية لا معنى له في اختبارات السرعة (Speeded tests) تعطي في زمن قصير) افرض ان لدينا اختباراً للعمليات الحسابية البسيطة مألفاً من ١٠٠ تمرين من نوع ٣ + ٥ = ؟، وان الاختبار يستخدم مع الكبار والوقت المحدد له دقيقتان سنحصل على فروق كبيرة في العلامات على الاختبار، ولكن هذه الفروق ستكون بشكل أساسي فروقاً في السرعة وستكون الأخطاء عاملا ثانوياً، والاحتمالات قوية ان الشخص الذي حصل على العلامة ٥٠ كان قد حاول الاجابة عن ٥٠ فقرة فقط، ٢٥ منها

فردية ٢٥ الأخرى زوجية، وبكلمة أخرى يظهر اتساق تام بين النصفين لأن الفرصة لمحاولة الاجابة عن الفقرات بشكل آلي تتعادل في نصفي الاختبار وقد ندر من الاختبارات ما اعتمد اعتماداً كلياً على السرعة، كما هو الأمر في المثال المذكور أعلاه، ومع ذلك فكثير منها يتضمن درجة من السرعة وعامل السرعة هذا يميل الى تضخيم الثبات المحسوب بطريقة التجزئة النصفية، وكمية الزيادة عن التقدير الحقيقي للثبات تعتمد على درجة السرعة في الاختبار وتعاظم الزيادة ـ غير الحقيقية ـ كلما تعاظم عامل السرعة في الاختبار.

ومع ذلك لابد أن يدخل عامل السرعة بدرجة كافية وعامة تبرر دوماً تجنب استخراج تقديرات الثبات بالتجزئة النصفية، وعلى مستخدمي الاختبارات أن يطلبوا من ناشري الاختبارات تزويدهم بتقديرات للثبات تستند الى الصور المتكافئة.

تقدير الثبات من احصائيات الفقرة:

ان المعلم أو الباحث الذي يكثر من اعتماده على الاختبارات ويقرأ البيانات التفصيلية التي كتبت عنها في أدلتها المنشورة لابد أن يصادف طريقة أخرى لتقديرات ثبات الاختبار من تطبيقه مرة واحدة فقط، وتعتمد هذه الطريقة على الاتساق في أداء الفرد من فقرة الى أخرى وتستند الى الانحراف المعياري للاختبار والانحرافات المعيارية للفقرات المفردة، وفي صورتها العامة يطلق عليها معامل الفا (Coefficient alpha) ويعبر عنها بالمعادلة التالية:

$$\text{الفا} = \frac{(\text{د})}{(\text{د} - ١)} \quad \frac{(\text{ع}^{٢}\text{ت} - \text{مج ع}٢)}{\text{ع}^{٢}\text{ت}}$$

حيث الفا: قيمة الثبات المقدرة.

د : عدد الفقرات في الاختبار.

ع ت: الانحراف المعياري للاختبار.

مج: تعني «خذ مجموع» وتشمل د من الفقرات.

ع د: الانحراف المعياري لفقرة.

وعندما تعطي كل فقرة العلامة (١) أو (صفر) وتناظر (ناجح) أو (راسب) على الفقرة فان: $\text{د}^{٢} = \text{ط طَ}$

حيث ط: نسبة الافراد الذين نجحوا في الفقرة.

و طَ: نسبة الأفراد الذين رسبوا في الفقرة.

عندئذ، تأخذ المعادلة (٢) الشكل التالي:

$$\text{ر}١١ = \frac{(\text{د})}{\text{د} - ١} \quad \frac{(\text{ع}^{٢}\text{ت} - \text{مج ط طَ})}{\text{ع}^{٢}\text{ت}}$$

وتسمى هذه المعادلة معادلة (كودر ـ ريتسارد سن رقم ٢٠) (Kuder-Richardson Formula # 20) نسبة الى واضعيها، ويشير الى الرقم ٢٠ الى ترقيم المعادلة في مقالتهما الأصلية، وبواسطتها يمكن الحصول على تقدير لما يسمى بالاتساق الداخلي (Internal Consistency) للاختبار ويشير الى الدرجة التي تشترك بها جميع فقرات

الاختبار في قياس خاصية معينة في الشخص، وعندما يكون الاختبار متجانساً (homogeneous) بمعنى ان كل فقرة فيه تقيس نفس العوامل العامة في القدرة أو الشخصية كما تقيسها أية فقرة أخرى، فان معامل الفا ومعادلة (كودر ـ ريتشارد سن) تعطيان تقديرات لها نفس المعنى الذي تعطيه التجزئة الى فقرات فردية وأخرى زوجية التي سبق عرضها أعلاه، ولذلك فان التقدير المسحوب بمعادلة (كودر ـ ريتشارد سن)، (١) لايأخذ في الحسبان التباين في الفرد من وقت الى آخر (٢) غير مناسب في اختبارات السرعة، فاذا أخذنا في الاعتبار هذين المحددين فانه يعطينا تقديراً محتفظاً (أي يميل أن يكون أقل بعض الشيء في القيمة العقلية) للثبات بطريقة التجزئة النصفية.

مقارنة بين طرق الثبات:

يلخص الجدول (٣) مقارنة بين الطرق المختلفة لتقدير الثبات، ويبرز أربعة عوامل يمكن أن تجعل من العلاقة على اختبار تقديرا غير دقيق لاداء الفرد الحقيقي، ويبين الجدول أي مصادر الخطأ ينعكس في كل أسلوب من الأساليب المذكورة آنفاً لتقدير الثبات، وعلى العموم كلما زاد عدد اشارات(X)في العمود، كلما كانت القيمة التي يحصل عليها لتقدير الثبات أكثر تحفظاً (أي أقل) ومن هنا يمكن ملاحظة كيف ان هذه الطرق ليست متكافئة، والطريقة التي تتيح لتأثيرات جميع مصادر التباين بالظهور هي فقط تطبيق لصور متكافئة للاختبار بفاصل زمني بينها.

الجـــــدول (٣)

مصادر التباين المتمثلة في الطرق المختلفة لتقدير الثبات

مصدر التباين	الاعادة مباشرة لنفس الاختبار	الاعادة بعد فترة لنفس الاختبار	الصور المتكافئة بدون فاصل زمني بينهما	الصور المتكافئة بفاصل زمني بينهما	التجزئة الى نصفين فردي زوجي	معامل الفا لاختبار مفرد
ـ كم يتوقع أن تتذبذب العلامة بسبب:						
ـ التباين الناشيء في اداة القياس	×	×	×	×	×	×
ـ التغير في الفرد من يوم لآخر		×		×		
ـ التغير في عينة المهمات الخاصة			×	×	×	×
ـ التغير في سرعة الفرد في الأداء	×	×	×	×		

أما الطرق الأخرى فكل منها يحجب مصدراً أكثر للتباين يمكن أن تكون له دلالته عند الاستعمال الفعلي للاختبار، واعادة تطبيق نفس الاختبار تغفل التباين الناشيء في عينة الفقرات، وفي أي أسلوب تتم عملية الاختبار فيه بنفس الوقت يغفل التباين في الفرد من يوم الى آخر.

أما عندما تتم عملية الاختبار كوحدة في حدود زمنه مفرده، تغفل السرعة في الاستجابة، ومن هنا فالحقائق التي يبرزها هذا الجدول يجب أن تؤخذ بعين الاعتبار في تقويم بيانات الثبات التي نجدها في أدلة الاختبارات المنشورة أو تقارير البحوث.

كم يجب أن يكون معامل الثبات مرتفعاً؟

من الواضح أنه: اذا تساوت اعتبارات أخرى معينة فاننا تفضل اداة القياس الأكثر ثباتاً ولكن كثيراً ما نتساءل: ما الحد الأدنى للثبات حتى يعتبر مقبولا؟ في الحقيقة لا توجد إجابة عامة عن هذا السؤال، واذا توجب علينا أن نتخذ قراراً أو اجراءاً ما يتعلق بفرد معين فما علينا الاَّ أن نفعل ذلك معتمدين على أفضل ما يتيسر لنا من معلومات مهما كان ثباتها، بشرط واحد وهو ان يكون الثبات أكثر من الصفر، (طبعاً هنا وفي جميع الأحوال، يكون الاعتبار الأول بصدق المقياس) ان تقويمنا لأي أداة جديدة يجب أن يكون دائما بمقارنتها بادوات اخرى تتنافس معها فاذا كان لدينا اختبار للرياضيات في المرحلة الثانوية، بمعامل الثبات ٠,٨٠ فانه يصبح غير مفضل اذا

توفرت لنا اختبارات أخرى يتراوح ثباتها بين ٨٥, و ٩٠, ومن ناحية أخرى فان أداة التقدير «القيادية» بمعامل ثبات لا يزيد عن ٦٠, يمكن أن يكون المفضل لدينا اذا كان البديل تقديرات مضبوطة ثباتها في حدود ٤٥, الى ٥٠, .

ومع أننا لا نستطيع ان نضع حداً أدنى مطلقاً لما يجب ان يكون ثبات المقياس، لكننا نستطيع ان نشير الى مستوى الثبات اللازم لتحقيق مستوى معين من الدقة في وصف فرد أو مجموعة، لنفرض أننا أعطينا اختباراً معيناً لشخصين وأن أحدهما (أ) وقع في المئين ٧٥ في مجموعته بينما وقع الشخص الثاني (ب) في المئين ٥٠ فما هو احتمال أن يظل (أ) متفوقاً على (ب) عند اختبارهما للمرة الثانية.

يبين الجدول (٤) هذا الاحتمال لقيم مختلفة من معاملات الثبات، فعندما يكون قيمة الارتباط صفراً فان الفرص هي بالضبط ٥٠ : ٥٠ ان ترتيب الشخصين سوف ينعكس، وعندما يصبح معامل الارتباط ٥٠, فاحتمال ان ينعكس الترتيب يزيد عن ١ : ٣، فاذا بلغ الارتباط ٩٠, فان فرص تبادل المواقع بين الشخصين عند تكرار الاختبار تصبح ١ من ١٢ ولكي نحصل على ٤ فرص من كل خمس فرص بأن يبقى الفرق في نفس الاتجاه نحتاج الى ثبات بقيمة ٨٠, .

الجـــــدول (٤)

النسبة المئوية لعدد المرات التي ينعكس فيها اتجاه الفرق عند اعادة الاختبار بين علامتين احدهما في المئين ٧٥ والثانية في المئين ٥٠

معامل الثبات	النسبة المئوية للمرات التي ينعكس فيها الفرق عند اعادة الاختبار		
	علامات افراد (كل على حدة)	متوسط مجموعات يتألف كل منها من ٢٥ فرداً	متوسط مجموعات كل منها ١٠٠ فرد
٠,٠٠	٥٠,٠	٥٠,٠	٥٠,٠
٠,٤٠	٤٠,٣	١٠,٩	٠,٧
٠,٥٠	٣٦,٨	٤,٦	٠,٠٤
٠,٦٠	٣٢,٥	١,٢	
٠,٧٠	٢٧,١	٠,١	
٠,٨٠	١٩,٧		
٠,٩٠	٨,٧		
٠,٩٥	٢,٢		
٠,٩٨	٠,٠٥		

ثبات الفروق بين العلامات:

في بعض الأحيان لا يكون اهتمامنا بالعلامات الافرادية بقدر ما هو بالعلاقة بين العلامات مأخوذة أزواجاً أزواجاً، ومن هنا يمكن

أن يكون اهتمامنا بالفروق بين الاستعداد المدرسي والتحصيل القرائي عند فريق من الطلبة، أو قد نهتم بدراسة التحسن في القراءة مقاساً بالفرق في الأداء على اختبار مبدئي أعطى الآن وآخر أعطي بعد أربعة أشهر، ان الأمر الهام في هذه الأمثلة بالنسبة لكل فرد هو الفرق بين العلامتين ولابد من أن نتساءل عن ثبات تقديراتنا لهذه الفروق اذا عرفنا خصائص الاختبارين اللذين استخلصت الفروق على أساسهما.

إن ثبات الفروق في الأداء على اختبارين هو في الواقع أقل بشكل ملموس من ثبات أي من الاختبارين كل على حدة، ويعلل ذلك على أساس عاملين اثنين.

بعد هذا الاستعراض السريع للصفات المرغوب توفرها في أي مقياس سواء كان تربوياً أو نفسياً فاننا ننتقل الى بحث صدق وثبات المقاييس التي تستخدم مع المجرمين سواء للتعرف على حالتهم الذكائية أو النفسية أو للكشف عن الجرائم عندهم، ان هناك بعض المقاييس التي تستخدم للتعرف على المستوى العقلي للمجرم وهذه المقاييس تأتي في مقدمتها مقاييس الذكاء، ومع أن للذكاء مقاييس متعددة بعضها فردي والبعض الآخر جمعي، وبعضها لغوي والآخر ادائي، وبعضها ثقافي والبعض الآخر غير ثقافي أو غير معتمد على الثقافة، الاَّ أنها جميعها تقوم على أساس أن الذكاء يتألف من مجموعة من العوامل يأتي في مقدمتها القدرة اللغوية، والقدرة العددية والقدرة على المحاكمة العقلية والقدرة على الحفظ والاسترجاع والقدرة على

حل المشكلات والقدرة على الادراك المكاني والقدرة على الميكانيكية . . وغيرها، وقد افترض (جيلفورد) وجود هيكل أو بناء للقدرات العقلية قوامه (١٢٠) قدرة أو عامل تم لحد الآن التعرف على ما يقرب من (٨٠) عاملا منها.

إن صدق هذه المقاييس هو من نوع صدق المفهوم، فمع أنه لا يوجد اتفاق تام على طبيعة الذكاء وماهيته الاَّ أنه يوجد اتفاق كبير على بعض المظاهر أو القدرات التي اذا تمثلت في الفرد قيل انه ذكي، وقد تم التأكد من صدق هذه المقاييس بعدة طرق وكيفيات، منها قدرتها على التنبؤ بالنجاح المدرسي، وعن طريق استخدام مبادىء التحليل العاملي أو عن طريق الترابط مع مقاييس أخرى، أو عن طريق التميز في ذكاء الأفراد من عمر لآخر على اعتبار أن الذكاء ينمو من العمر وبخاصة في سنوات العمر العشرين الأولى.

أما من ناحية الثبات فان هناك عدة طرق تم بموجبها التأكد من الثبات يأتي في مقدمتها اعادة الاختبار أو استخدام صور متكافئة له حيث أن بعض هذه الفحوص يأتي على أكثر من صورة.

ومن الجدير بالذكر أن من أشهر فحوص الذكاء المستخدمة على نطاق واسع في العالم هي مقاييس: (ستانفورد ـ بنية ويكسلر، لورج، ثورندايك رافن) وغيرهم، وقد تم نقل الكثير منها الى العربية وهي لا زالت تستخدم على نطاق ضيق ولغايات البحث فقط، ان مقاييس الذكاء على اختلاف أنواعها تتمتع بدرجات عالية من الصدق

والثبات وهي أدوات صالحة تماماً للأغراض التي وضعت من أجلها وتستعمل على نطاق واسع.

أما الاختبارات التي تستخدم في الكشف عن المجرمين فهي محدودة العدد نوعاً ما ويأتي في مقدمتها:

١ ـ اختبار رورشارخ.
٢ ـ اختبار تفهم الصور.
٣ ـ اختبار التداعي الحر.
٤ ـ زمن التفاعل العقلي.
٥ ـ قياس الانفعالات النفسية عن طريق الجلد.
٦ ـ الجهاز ذو التسجيلات المتعددة البوليغراف.

١ ـ اختبار رورشارخ:

أما اختبارات بقع الحبر فقد نشأت في بادىء الأمر في ألمانيا منذ عام ١٨٥٧ ولكنها أخذت أهميتها في عام ١٩٢١ على يد عالم النفس السويسري (هيرمان رورشارخ) حيث نشر بقع الحبر التي استخدمها، بالاضافة الى النتائج الاكلينيكية التي توصل اليها والفلسفة النظرية التي تقوم عليها هذه البقع، وهذه البقع من الواجب حسب رأي (رورشارخ) أن تتميز بالصفات التالية: البساطة النسبية، التماثل، امكانية الايحاء، وقد توصل (رورشارخ) الى بقع الحبر التي تعرف باسمه بعد ابحاث ودراسات مضنية في المستشفيات النفسية استمرت ١٠ سنوات، ويعتقد (رورشارخ) أنه من الممكن التعرف على الصفات الأساسية لشخصية الفرد من خلال استجاباته

للبقع التي تعرض عليه، ويؤكد الفرد (بنية) العالم الفرنسي أن بقع الحبر يمكن استخدامها في الكشف عن القدرة على التخيل البصري وذلك تمهيداً لدراسة عدد من الصفات الخاصة بشخصيات الأفراد.

ان طريقة (رورشارخ)قائمة على نظرية عامة للشخصية وعلى الارتباطات القائمة بين مختلف أجزائها حيث أن اجابات المفحوصين على هذه البقع عندما يتم تصنيفها في مجموعات رسمية فانه يمكن من خلالها التعرف بطريقة موضوعية على شخصيات الأفراد بوجه عام، لقد ساعدت بقع الحبر في التعرف على نفسيات الأفراد غير الأسوياء كالأحداث المنحرفين، ومدمني الكحول والمصابين باللعثمة والمصابين بمرض الانفصام وما شابه.

ان الافتراض الأساسي الذي يقوم عليه أسلوب بقع الحبر لرورشارخ هو وجود علاقة بين الشخصية والادراك، ان الطريقة التي يدرك بها الفرد بقع الحبر تعكس مظاهر أساسية في تكوينه النفسي، ان بقع الحبر هي مثيرات مناسبة في هذا المجال لأنها مثيرات غامضة، فهي لا تعكس استجابات متعلمة ولكنها تترك المجال لظهور عدد متنوع من الاستجابات الممكنة، فعندما يطلب الى الفرد أن يذكر ما يراه في بقعة الحبر فعليه أن يرد بطريقة غير متعلمة وشخصيته، ان ادراكاته يتم اختبارها وتنظيمها بدلالة حاجاته وخبراته وعاداته في الاستجابة بالاضافة الى الصفات الفيزيقية للبقع نفسها.

وفي ضوء الافتراضات السابقة فان تحليل استجابات أي فرد من الأفراد لبقع الحبر يجب ان تعكس أشياء مثل: طبيعة الأمور

الداخلية التي تحرك سلوكه وعن دوافعه وبواعثه وقدرته على الانضباط وطريقة مهاجمته للمشكلات التي تواجهه، بالاضافة الى بعض مميزاته الشخصية الأخرى، انها وسيلة تشجع الفرد في أن يبوح بعالمه الداخلي الذي يخفيه في العادة عن الآخرين، ان الشخص في استجابته لمواقف الحياة العادية غالباً ما يخفي دوافعه الحقيقية، ولكن في حال بقع الحبر فانه يستجيب لأشياء غامضة لا يعرف ما هي الاجابة المقبولة أو غير المقبولة، ولذلك فانه يستجيب دون محاولة للتزييف، انه يستجيب بطريقة تلقائية ولذلك فانه وبطريقة لا شعورية يكشف عن خفايا نفسه، انه يكشف عن السلوك أو الشعور غير المتعلم.

ان ما يتصل بصدق أسلوب (رورشارخ) أمر يصعب التحقق منه بالطرق السابقة، ان هذا الأسلوب يقوم على نظرية معينة مفادها ان الانسان في ادراكه للمواقف الغامضة يحاول أن يكملها من ذاته ويحاول أن يسقط عليها من مميزاته الشخصية، وقد تم التأكد من صحة هذا الافتراض بشكل جزئي في استخدام المقياس في حالات مرضية معينة. أي أن المقياس يتمتع بصدق اكلينيكي أكثر من تمتعه بصدق احصائي.

وبالطبع فان هذا الأمر لا يقلل من شأن المقياس ولا يضعف من استخداماته لانه في العادة لا يستعمل لوحده كأداة للكشف عن نفسيات الأفراد وانما تقترن نتائجه عادة بنتائج مقاييس أخرى مكملة له، وبالنسبة للثبات فهو أيضاً من مستوى متدن احصائيا شأنه في

ذلك شأن الصدق، وهو قلما يصل الى (٠,٤٠) على أحسن درجة، وفي هذا المجال قد يفيد أن نذكر بأن علماء النفس ينظرون الى هذه النوعية من المقاييس على أنها مفيدة وانها أفضل من لا شيء انه من المعروف في الأساط النفسية ان اختبار (رورشارخ) يفضل دائما أن تقترن نتائجه مع نتائج تفهم الموضوع الذي سيلي شرحه لأن كلا منهما يكمل الأخر.

٢ ـ اختبار تفهم الموضوع:

ان اختبار تفهم الموضوع (TAT) هو عبارة عن مجموعة من الصور التي تمثل مواقف غامضة ولكنها أقل غموضاً من بقع الحبر وفي كل حالة يطلب الى المفحوص ان يروي قصته عما يراه في الصورة ومن المفترض أن المفحوص على محتوى هذه القصص يسقط خبراته الماضية واحتياجاته المستقبلية التي لا يمكن الافصاح عنها في الظروف العادية، ويستدل على هذه الصفات عادة اذا كان هناك تكرار لنمط معين من الاستجابات من قصة لأخرى أو في حالة عدد من هذه القصص.

ان فحص تفهم الموضوع مفيد في أية دراسة متكاملة للشخصية وفي تفسير الاضطرابات السلوكية والأمراض النفسية والعصبية والذهنية، وهو مفيد كفحص أولي قبل البدأ بأية دراسة من نوع التحليل النفسي، ان المفحوص يترك موقف الفحص وهو غير مدرك بأنه قد أعطى ما يشبه الصورة الشعاعية لذاته الداخلية.

ان ما ذكر عن الصدق والثبات في حالة اختبار بقع الحبر يصدق هنا أيضاً لأن الاختبارين يقومان على نفس المبادىء السيكولوجية مع أن مراتبهما في هذه الحالة أعلى بعض الشيء، وعلى كل حال فقد أظهر اختبار تفهم الموضوع، كما أظهر اختبار بقع الحبر فائدة اكلينيكية ملحوظة الأمر الذي جعل من انتشار استخدامها أمراً مقبولاً.

٣ ـ اختبار التداعي اللفظي:

فيما يتصل بظاهرة ارتباط الأفكار من الوجهة النظرية توجد ثلاثة عوامل رئيسية تؤثر في متانة «الروابط الفكرية» وهي الجدة أو حداثة العهد، والتكرار، والأهمية ويلحق بالأهمية الغرابة والشذوذ والروعة والانفعالات النفسية وبالجملة كل ما يثير في النفس أي اهتمام، فاذا ارتكب شخص جريمة خطيرة كجريمة قتل مثلا، فان هذا الحادث يترك أثراً شديداً يربط الوقائع في ذهنه ربطاً محكماً يدوم زمناً طويلا، وبالأخص اذا كان المتهم مجرماً بالمصادفة أو حديث العهد بالاجرام لم تجمد عواطفه ولم تنضب موارد تأثيراته وانفعالاته فان عقله يصبح شديد التنبه بكل ما له علاقة بجرمه، عظيم القابلية للتأثر بذكرياته وتضحى نفسيته ميداناً صالحاً يباشر فيه الباحث تجاربه واختباراته واذا كان الجرم قريب الوقوع ولم يمض عليه زمن تكون قد هدأت فيه ثورة نفس مقترفة اجتمع بذلك لدى الخبير عاملان قويان يعملان على تقوية الخواطر التي ارتبطت في ذهن المتهم بوقائع الجريمة ويساعدان على تشخيص نفسيته تشخيصاً فنياً دقيقاً وهما عاملا الخطورة والجدة.

فاذا فرض وكان غرضنا اختبار شخص متهم بجريمة قتل مثلاً لنعرف أهو الذي ارتكبها أم لا، وكان المتهم منكراً معرفته القتيل انكاراً باتاً، ويدعي أنه لم يره ولم يدخل منزله في حياته قط، ولا يعرف عن مكان الحادث ووقائعه وظروفه شيئاً فاننا اذا عرضنا عليه أمر اختباره «بالتداعي اللفظي» (The Word Association Test) فهو لا يمانع عادة اجراء هذا الاختبار لانه يبدو له تافهاً لأول وهلة.

اما أدوات الاختبار فتنحصر في «كرونسكوب» أو ساعة خاصة تسمى ساعة «ضبط» (Stop watch) تقدر الفترات الزمنية بعشر الثانية، وقائمة مجهزة من عدد وافر من الكلمات المختلفة (وجرت العادة ان تكون مائة كلمة) ينتخب من بينها ثلاثون أو أربعون لها صلة بالجريمة ووقائعها وظروفها ومحتوياتها والاجراءات والتدابير التي اتخذت لتنفيذها والأشخاص الذين اشتركوا فيها أو كانوا حاضرين وقت وقوعها والبواعث التي أدت الى ارتكابها وهلم جرا.

أما طريقة الاختبار فهي ان يطلب من المختبر أن يلفظ أول كلمة ترد على ذهنه بمجرد سماعه كل كلمة تلقى عليه من الكلمات الواردة في القائمة على أن يكون جوابه عليها بأقصى ما يمكن من السرعة، ثم ترصد الأجوبة في مذكرة كما ترصد أزمنتها بدقة بواسطة (الكرونسكوب) أو الساعة واعداد القائمة المئوية هو أول عمل تجب العناية به عناية خاصة، لأن اختبار الكلمات التي لها ارتباط بالحادث وكيفية توزيعها بين باقي الكلمات الواردة بالقائمة مما يتطلب مهارة فنية وخبرة واسعة، كما أن كشف التأثرات النفسية وتحديد مدلولها

يحتاج لكثير من دقة الفكر وقوة الملاحظة، اذ قد يتفق ان يكون للمتهم من القدرة وقوة ضبط النفس ما يكفي لعدم اظهار أفكاره وخواطره الحقيقية مباشرة على أنه مع ذلك يمكن للباحث المدقق كشفها بطريقة استدلالية أو غير مباشرة لأنه بسبب وقوع مثل هذه الخواطر تحت ضغط الارادة تبقي ماثلة في المخيلة، وكثيراً ما يظهر أثرها في الجواب التالي أو الذي يليه وربما انطلقت بذاتها في خلال الاستجواب مدفوعة بقوة ضغط الانفعال المحتبس.

فبدراسة طبيعة مثل هذه الردود المترتبة على الأسئلة المتعددة المتصلة بالحادث والموجهة بدقة واحكام في أثناء الاختبار، قد يوفق المرء الى استخلاص وصف مسهب وصورة واضحة لمكان الحادث وكيفية ارتكابه واسماء من عاونوا المتهم فيه وغير ذلك من التفصيلات وهو في غفلة عما يبدو منه ويشهد به قلبه ولسانه.

٤ ـ زمن التفاعل العقلي:

والفائدة من قياس سرعته في البحث الجنائي ان للانفعالات النفسية عند اثارتها وقت الاختبار تأثيراً في تأخير سرعة الخواطر بسبب الصدمة التي تحدثها الذكريات المثيرة حال ايقاظها يظهر في الكرونسكوب بوضوح فبقياس سرعة التداعي بهذا الجهاز الدقيق وتقدير زمنها يمكن كشف آثار الانفعالات النفسية حال وقوعها ورصدها بدقة، وبالتالي يمكن تعيين الكلمات التي أثارت هذه الانفعالات وتقدير أوجه ارتباطها بعقلية المتهم من جهة، ثم بوقائع

الجريمة من جهة أخرى. وبذلك يمكننا تحديد موقف المتهم بين الادانة والبراءة في حادث معين، واظهار المدمن من البريء من بين عدة أشخاص وقعت عليهم الشبهة، فرصد زمن التداعي وملاحظة ما يطرأ عليه من السرعة والبطء قد يفوق في أهميته كل ما تقدم، فان الكلمات التي تثير انفعالا نفسياً من شأنها أن تعوق النشاط العقلي بدرجة محسوسة بحيث لا ترد تلبيتها في أقل من ضعف أو ثلاثة أضعاف أمثال الزمن العادي حتى رأى بعضهم أنه لا حاجة الى قياس الزمن بجزء من الف من الثانية في الأبحاث الجنائية بل يكفي لذلك أن يقدر بجزء من مائة أو عشرة وأحياناً بجزء من الثانية وسنقتصر في بحثنا الحالي على جعل عشر الثانية وحدة زمنية في قياس سرعة التفاعلات العقلية واظهار ما بينها من الفروق لأن في ذلك الكفاية، ويمكننا أن نستخدم لهذه الغاية عوضاً عن الكرونسكوب ساعة خاصة لها عقرب كبير يقطع دورته حول وجه المقسم الى (٣٠) درجة في ثلاث ثوان، ويمكن تحريكه أو أيقافه بالضبط على مقبض الساعة، فعند ذكر كلمة التنبيه يضغط الانسان على المقبض فيدور العقرب، وبمجرد ذكر كلمة التنبيه يضغط عليه ثانية فيقف وبذلك يمكن قياس الزمن بعشر الثانية.

وقد عرفنا مما سلف ان الزمن اللازم لاستحضار الأفكار أو زمن التداعي يختلف باختلاف الأفراد فان الأشخاص البطيئي الفهم ربما يحتاجون لضعف الزمن الذي يحتاج اليه سريعو الفهم لايقاظ الخواطر في أذهانهم وتنبيهها، ولذا فان أول ما يتبين على المختبر عمله هو استخراج متوسط سرعة الخواطر العادية للشخص الموضوع تحت

الاختبار وهو ما يسمى «بالمعادلة الشخصية» (Personal Equation) وبعد ذلك يوجه عنايته واهتمامه الى ما يطرأ على مستوى هذه السرعة من التغيير فاذا كان متوسط السرعة في الأحوال العادية لا يزيد على نصف ثانية لشخص متصف بحدة الذهن والذكاء، ثم شوهد خلال الاستجواب ان استحضار بعض الخواطر استغرق منه ثانية ونصفا أو ثانيتين، أعني ثلاثة أو أربعة أمثال الزمن العادي فان ذلك يكون داعياً للشبهة.

وقد يتفق أحياناً ان الصدمة النفسية في احدى التداعيات تكون من الشدة بحيث يتعدى أثرها الى التداعي التالي أو الذي يليه أو الذي بعده، حتى ولو كانت الكلمات في كل منها عادية لا علاقة لها بالجريمة لأن قوة الصدمة من شأنها أن تحدث ارتجاجاً في سلسلة الخواطر، وتؤثر في النشاط العقلي تأثيراً محسوساً، يظهر من تأخير زمن التفاعل في مجموعة من التداعيات، وقد يتم التداعي المحرج في زمنه الطبيعي ولا يظهر التأخير الا فيما يليه من التداعيات بسبب حدوث الصدمة على أثر النطق بالجواب المحرج.

فتحليل التأخيرات الزمنية ودرس طبيعة الخواطر التي سببت هذه التأخيرات هو أهم عمل في اجراءات الاختبار يجب على الباحث ان يوجه اليه معظم اهتمامه ويشمله في عنايته الخاصة.

٥ ـ قياس الانفعالات النفسية عن طريق الجلد:

ان كشف آثار الانفعالات النفسية لايقف عند حد قياس زمن التفاعلات العقلية بالكرونسكوب بل هناك من الوسائل العلمية

الأخرى ما لا يقل عنه خطورة وأهمية، انه من الممكن الكشف عن التأثيرات النفسية بقياس قوة مقاومة سطح الجلد للتيارات الكهربائية بتمرير تيار كهربي ضعيف في الجسم، ثم رصد آثار الانفعالات التي تنتاب الشخص الذي تحت الاختبار بمقياس دقيق لمقاومة التيار معروف باسم «جلفانومتر» (وهو جهاز ذو إبرة مغناطيسية حساسة للتيارات الكهربية وتسجل ما يطرأ عليها من التغيير) فان هذه الظاهرة الجلدية العجيبة قد درست من الوجهة الطبية وأجريت عليها عدة أبحاث فنية قام بها نخبة من خيرة العلماء أمثال شارل فريه (Feree) وفيجوروكس (Vigouroux) وستكر (Sticker) وتارشنوف (Tarchanof) وزمر (Sommer) وفورستنو (Firstenau) وغيرهم من رجال العلم فدلت التجارب المتعددة التي قاموا بها على أن كل خاطر يمر بالبال أو كل لفظ يطرق السمع مهما كان مدلوله تافهاً وظاهره مجرداً عن المغزى له تأثيره الخاص في النفس ويظهر بوضوح في جهاز الاختبار.

أما أساليب الاختبار فكثيرة متنوعة ويكفي هنا ذكر الطريقة الأكثر شيوعاً وهي ان يؤمر المختبر بأن يستوى على مقعد مريح وثير الفراش ذي متكأين يركز عليهما منكبيه ثم يبسط راحتيه على لوحين نحاسيين أو بعض أطراف أصابعه في اناءين صغيرين في الصيني، بهما محلول ملح الطعام، ثم يسلط على اللوحين أو الاناءين تيار كهربي ضعيف من بطارية قوتها نحو (١٥ فولتاً) بحيث لا يكون التيار محسوساً طول مدة الاختبار وتقاس التقلبات التي تطرأ على التيار

«بجلفانومتر» الدقيق ذي مرآة عاكسة ويوضع على مسافة متر تقريباً تجاه المرآة مصباح قوي يلقي أشعته عليها وهذه تعكس على لوح كبير من السليوليد، قسم الى سنتيمترات وملليمترات بحيث ان كل تغيير في قوة مقاومة التيار مهما كان طفيفاً يؤدي الى تحرك ابرام الجلفانومتر بالمرآة فيظهر أثر هذه الحركة مكبراً في رحلات الأشعة الساقطة على اللوح المقسم، وبهذا يمكن قياس درجات انحراف الأشعة عن موضعها الأصلي بسهولة فكل لفظ يحدث في النفس انفعالا مهما كان طفيفاً في مقاومة الجلد للتيار له تأثير خاص يبدو على اللوح مكبراً وبذلك يتسنى للخبير أن يرصد مراتب الانفعالات المختلفة لكل خاطر بدقة ووضوح تام.

٦ ـ كشف الجرائم وآثار الانفعالات النفسية:

آثار الانفعالات النفسية يمكن تعريفها بأنها هي الأعراض التي تبدو على الأعضاء الظاهرة أو الباطنة لجسم الانسان بسبب العوامل التي تؤثر في تلك الأعضاء تأثيراً خاصاً كالسرور والحزن والغضب والخوف والتهيج (Excitement) والهبوط وغيرها، فكل حالة من هذه الحالات لها تأثيرها الخاص في المجموع العصبي وبجملته جهاز الحركة

وقد دلت الحركة على أن حالتي السرور والحزن وهما اجمالا الحالتان الرئيسيتان اللتان تتفرع عنهما معظم الحالات الأخرى لهما نتيجتان عكسيتان من حيث التأثير في أعضاء الجسم واعراضهما

متضادة، فحالة السرور تزيد في مدى الحركات الخارجية وتجعلها أكثر من المعتاد وتقلل من مدى الحركات الداخلية وتجعلها أقل من المعتاد، بخلاف حالة الحزن فانها تؤدي الى عكس ذلك أي الى أضعاف الحركات الخارجية وتقوية الحركات الداخلية بمعنى أن حركة الانبساط تقل وحركة الانقباض تزيد أكثر من المألوف.

٧ ـ الجهاز ذو التسجيلات المتعددة ـ البوليغراف أو الذتم

البوليغراف كلمة اصطلاحية معناها حرفياً «ذو التسجيلات المتعددة» ويمكن اختصارها بكلمة واحدة مؤلفة من الأحرف الأولى لهذه الكلمات الثلاث أعني (الذتم) للدلالة على جهاز مركب من جهازين فأكثر من الأجهزة المعدة لرصد آثار الانفعالات النفسية في الأعضاء المتمتعة بالحركة الذاتية التي مر بنا ذكرها: كاللبنوموغراف (أعني رسام حركات التنفس) والفجموغراف (رسام حركات النبض) والبلتزموغراف (رسام حركات توارد الدم) والسيكوجلفانومتر (مقياس انفعالات الجلد الكهربائية) والكرنسكوب (كاشف الزمن) وما اليها.

ويسمى البوليغراف أحياناً (Lie detector) أعني كاشف الأكاذيب أو بعبارة موجزة «المفضاح» اشارة الى كونه يفضح ما تخفيه النفوس وما تطوي عليه السرائر من أكاذيب ومخاتلات أو غش أو خداع.

وطريقة استعماله تتلخص في اجلاس الشخص المراد اختباره على كرسي وثير ذي متكأين ثم يركب كل جهاز فرعي من أجهزة التسجيل في موضعه على العضو المخصص لرصد انفعالاته، وبعد احكام تركيب الأجهزة والتأكد من صلاحية كل منها لأداء مهمته يقوم الخبير بتوجيه أسئلة الى الشخص الموضوع تحت الاختبار، بعضها يدور حول أمور عادية لا احراج فيها والبعض يدور حول الجريمة أو الوقائع المحرجة التي يرد كشفها ومعرفة مدى صلة المختبر بها، فكل محاولة لاخفاء الحقيقة أو تغيير الأجوبة بقصد التضليل يظهر أثرها على الشريط المتحرك الذي نسجل عليه موجات الانفعال وذبذبات بواسطة مؤشرات التسجيل، فعندما يكذب المختبر يتميز شكل الموجات من حيث ارتفاعها وعمقها وعرضها، وتعرج خطوط حدودها، وما الى ذلك من الفروق التي تبدو جلية واضحة لكل ذي خبرة وكذا اذا ووجه المختبر بسؤال محرج أو طرقت سمعه عبارة أو كلمة أثارت من نفسه انفعالا.

ولعل من أكثر هذه الأجهزة ذيوعاً جهاز وضع تصميمه العلامة (رودلف كيلر) وقد سمي باسمه «بوليغراف كيلر» (Keeler Polygraph) وهو جهاز مركب من أربعة أجهزة لرصد حركات التنفس والنبض وتوارد الدم في الأطراف وانفعالات الجلد الكهربائية.

ومما لا ريب فيه أن رصد الانفعالات بالجهاز «ذي التسجيلات المتعددة» أو «الذتم» أقوى ضماناً وأسلم عاقبة من استخدام الأجهزة

الفردية، فان «الذتم» مثله مثل الطبيب الذي يعتمد في تشخيص الداء على مجموعة من وسائل الفحص الطبي والتحاليل عوضاً عن اكتفائه بوسيلة واحدة، اذ يكشف الطبيب من أعراض يكون أدنى الصواب في معرفة حقيقة العلة وتشخيص الداء، فان سرعة نبضات القلب قد تكون عرضاً مشتركاً بين مجموعة من الأمراض، ولكن اذا ما لجأ الطبيب بجانب ذلك الى قياس درجة حرارة المريض وضغط دمه وفحص كرياته وفحص صدره وما الى ذلك من وسائل الفحص كان له من ذلك أقوى ضمان لتشخيص سليم، وخير عاصم له عن الخطأ والزلل، كذا الحال فيما يختص بأعراض الكذب والمخاتلة حين تتضافر في كشفها مجموعة من الأجهزة المؤتلفة فانها تكون في دلالتها أدنى الى الصواب والدقة من أعراض يسجلها جهاز واحد.